TAIL PROBABILITY APPROXIMATION FOR ONE AND TWO SAMPLE U-STATISTICS BY THE SADDLEPOINT METHOD

By

DIVAKAR SHARMA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1993

Copyright © 1993

by

Divakar Sharma

To My Wife,
Supriya

ACKNOWLEDGEMENTS

I would like to acknowledge all those hidden contributors who have been a constant source of inspiration and shown unlimited support during my research. I wish to express my sincere appreciation and thanks

- to my advisor, Dr. Myron S. Chang, for his advice, guidance and constructive criticism and for providing a constant stimulus to improve the quality of research;
- to my favorite long-term companion, Supriya, for sharing with me all the pain and frustration experienced during this period, for being completely supportive and understanding and above all for being a constant source of encouragement;
- to Dr. M. Ghosh, Dr. P. V. Rao, Dr. J. Shuster and Dr. M. Rao for providing suggestions to further improve the manuscript;
- to my parents, Pran and Radha, for exposing me to all basic operations of human functioning and making me strong enough to survive; to my brother, Lalit, for reminding me why I should always strive to be better;
- to Gail and Laurie for making it possible to bring this manuscript into print;
- and finally to all my friends and colleagues in CBARDS at Merck Research Laboratories for the support I received in completing my research and degree.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		iii
LIST OF TABLES		vi
LIST OF FIGURES		x
ABSTRACT		xii
CHAPTERS		
1	INTRODUCTION	1
1.1	Review	1
1.2	Problem	2
2	SADDLEPOINT APPROXIMATIONS	6
3	ONE SAMPLE: U-STATISTIC	15
3.1	Notation and Results	16
3.2	U-Statistic: Degree K	26
3.3	Example	36
4	TWO SAMPLE: U-STATISTICS	49
4.1	Two Sample: Degree (1,1)	50
4.2	Two Sample: Degree (K_1,K_2)	64
4.3	Example	77
5	COMPARISON OF THE TAIL PROBABILITY APPROXIMATION BY THE METHOD OF EDGEWORTH EXPANSION AND SADDLEPOINT APPROXIMATION	88
6	POWER OF ONE SAMPLE U-STATISTICS	97
6.1	Discussion and Methods	97
6.2	Results	103

7 POWER: TWO SAMPLE U-STATISTICS 118

7.1 Discussion and Methods 118
7.2 Results 125

REFERENCES 138

BIOGRAPHICAL SKETCH 140

LIST OF TABLES

Page

3.1 Comparison of tail probability approximation of Wilcoxon signed rank test statistic (N=17) between exact (null) distribution, normal approximation and saddlepoint approximation 44

4.1 Comparison of tail probability approximation of Wilcoxon-Mann-Whitney test statistic (n=5,m=10) between exact (null) distribution, normal approximation and saddlepoint approximations 82

5.1 Comparison of tail probability approximation of Wilcoxon signed rank test statistic (N=17) between exact (null) distribution, normal approximation, Edgeworth expansion and saddlepoint approximation 93

6.1 Computation of expectations needed to approximate the cumulant generating function for power calculation of Wilcoxon Signed Rank test under normal shift alternatives 104

6.2 Computation of expectations needed to approximate the cumulant generating function for power calculation of Wilcoxon Signed Rank test under logistic shift alternatives 105

6.3 Computation of expectations needed to approximate the cumulant generating function for power calculation of Wilcoxon Signed Rank test under Cauchy shift alternatives 105

6.4 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under normal alternative (Δ=0.253347) between exact distribution, normal approximation and saddlepoint approximations 107

6.5 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under normal alternative (Δ=0.524401) between exact distribution, normal approximation and saddlepoint approximations 108

6.6 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under normal alternative (Δ=0.841621) between exact distribution, normal approximation and saddlepoint approximations 109

6.7 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under logistic alternative (Δ=0.405465) between exact distribution, normal approximation and saddlepoint approximations 110

6.8 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under logistic alternative (Δ=0.847298) between exact distribution, normal approximation and saddlepoint approximations 111

6.9 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under logistic alternative (Δ=1.38629) between exact distribution, normal approximation and saddlepoint approximations 112

6.10 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under Cauchy alternative (Δ=0.32492) between exact distribution, normal approximation and saddlepoint approximations 113

6.11 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under Cauchy alternative (Δ=0.726543) between exact distribution, normal approximation and saddlepoint approximations 114

6.12 Comparison of tail probability approximation of Wilcoxon Signed Rank test statistic (N=17) under Cauchy alternative (Δ=1.37638) between exact

distribution, normal approximation and saddlepoint approximations 115

7.1 Computation of expectations needed to approximate the cumulant generating function for normal shift alternatives 123

7.2 Computation of expectations needed to approximate the cumulant generating function for logistic shift alternatives 124

7.3 Computation of expectations needed to approximate the cumulant generating function for Cauchy shift alternatives 124

7.4 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under normal alternative (Δ=0.253347) between exact distribution, normal approximation and saddlepoint approximations 126

7.5 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under normal alternative (Δ=0.524401) between exact distribution, normal approximation and saddlepoint approximations 127

7.6 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under normal alternative (Δ=0.841621) between exact distribution, normal approximation and saddlepoint approximations 128

7.7 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under logistic alternative (Δ=0.405465) between exact distribution, normal approximation and saddlepoint approximations 129

7.8 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under logistic alternative (Δ=0.847298) between exact distribution, normal approximation and saddlepoint approximations 130

7.9 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under logistic alternative (Δ=1.38629) between exact distribution, normal approximation and saddlepoint approximations 131

7.10 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under Cauchy alternative (Δ=0.32492) between exact distribution, normal approximation and saddlepoint approximations 132

7.11 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under Cauchy alternative (Δ=0.726543) between exact distribution, normal approximation and saddlepoint approximations 133

7.12 Comparison of tail probability approximation for Wilcoxon-Mann-Whitney test statistic (n=5, m=10) under Cauchy alternative (Δ=1.37638) between exact distribution, normal approximation and saddlepoint approximations 134

LIST OF FIGURES

		Page
3.1	Wilcoxon Signed Rank test. Comparison of cumulant generating function using exact, cumulant generating function (approximate) and cumulant generating function (K4)	37
3.2(A)	Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	45
3.2(B)	Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	46
3.2(C)	Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	47
4.1	Mann-Whitney-Wilcoxon test. Comparison of cumulant generating function using exact, cumulant generating function (approximate) and cumulant generating function (K4)	79
4.2(A)	Mann-Whitney-Wilcoxon test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	84
4.2(B)	Mann-Whitney-Wilcoxon test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	85
4.2(C)	Mann-Whitney-Wilcoxon test. Comparison of tail probability approximation using exact, saddlepoint and normal distribution	86
5.1(A)	Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint (approximate cumulant generating function), normal and Edgeworth expansion	94

5.1(B) Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint (approximate cumulant generating function), normal and Edgeworth expansion 95

5.1(C) Wilcoxon Signed Rank test. Comparison of tail probability approximation using exact, saddlepoint (approximate cumulant generating function), normal and Edgeworth expansion 96

6.1 Wilcoxon Signed Rank test. Comparison of left tail probability approximation under normal alternative ($\Delta=0.841621$) using exact, saddlepoint and normal distribution 116

6.2 Wilcoxon Signed Rank test. Comparison of left tail probability approximation under Cauchy alternative ($\Delta=1.37638$) using exact, saddlepoint and normal distribution 117

7.1 Mann-Whitney-Wilcoxon test statistic. Comparison of left tail probability approximation under logistic alternative ($\Delta=0.847298$) using exact, saddlepoint and normal distribution 135

7.2 Mann-Whitney-Wilcoxon test statistic. Comparison of left tail probability approximation under Cauchy alternative ($\Delta=0.726543$) using exact, saddlepoint and normal distribution 136

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

TAIL PROBABILITY APPROXIMATION FOR ONE AND TWO SAMPLE U-STATISTICS BY SADDLEPOINT METHOD

By

Divakar Sharma

May 1993

Chairman: Myron Chang
Major Department: Statistics

This dissertation proposes an alternative to normal (large sample) approximation for the tail probability approximation (under both null and alternative hypotheses) for one and two sample U-statistics by saddlepoint methods.

Under minimal assumptions, approximations for the cumulant generating function of both one and two sample standardized U-statistics are derived by using Hoeffding's decomposition. These approximations are then used in Daniel's formulation (for a discrete case) of Lugannani and Rice's saddlepoint approximation to the tail probability.

This method has been applied to both the Wilcoxon Signed Rank test and the Mann-Whitney form of Wilcoxon's Rank Sum test to approximate the tail probability under null and various shift normal, logistic and Cauchy) alternatives. Finally, results using our approximation have been compared to both normal approximation and exact distribution (based on 10 million simulations) for both the examples.

CHAPTER 1
INTRODUCTION

1.1 Review

P. R. Halmos (1946) is credited with first creating interest in the existence of U-statistics. Wassily Hoeffding (1948) formally introduced them in his *Annals of Mathematical Statistics* paper. Since that time, many articles, along with quite a few books, have extended the theory and introduced new applications associated with U-statistics. The class of U-statistics is important for the following reasons:

- A number of commonly used statistics are, in fact, U-statistics.
- The simple but unified structure of U-statistics makes them ideal for studying general inference processes and for generalizing those parts of asymptotic theory concerned with the behavior of sequences of sample means.
- Application of the theory often generates new statistics useful in practical inference problems.
- Minimal assumptions regarding the underlying distributions are needed.

Even with all this attention focused on U-statistics, it has traditionally been hard to handle quite a few U-statistics based tests which have immense potential for numerous applications. The main difficulty has been the conspicuous absence of exact methods in general for the whole class.

1.2 Problem

Methods to estimate the tail probabilities (under both null and alternatives) exist for only a few U-statistics. Even for those U-statistics for which exact methods do exist, calculation of tail probabilities can be fairly time consuming, even for moderate sample sizes. Moreover, even if they exist, they may not be as well known as one might expect.

Normal approximations based on asymptotic distributions give an alternative approach for approximating tail probability, but are not very good if either the sample size is not large enough or the underlying distribution is heavy tailed. For power calculation, even for normal shift alternatives in one and two sample settings, if Δ (shift parameter) is not small enough, it is quite possible that normal (large sample) approximation may not be very good (Lehmann, 1975).

Simulation at times can solve the problem of estimating the tail probability for the U-statistics; however, it can obviously be influenced by the accuracy of random number generators. As a result of random fluctuations, there is always a chance of coming up with strange results for a tail probability.

The main goal of this research was to come up with a tail probability approximation via saddlepoint methods for the general class of U-statistics, one sample degree K and two sample degree (K_1,K_2), which could be used under both null and alternative hypotheses. Numerical examples (both in one and two sample settings) show that the proposed method is better than normal approximation and other methods such as Edgeworth expansion.

Lugannani and Rice (1980) derived a saddlepoint approximation for sum of n independent and identically distributed random variables. According to I. M. Skovgaard (1987), the approximation can be used for an arbitrary sample size n, but they note that it is applicable for n = 1. We decided to apply Lugannani and Rice's (1980) saddlepoint approximation for the distribution function under a discrete distribution for the standardized U-statistics, i.e., n = 1. H. E. Daniels (1987) derives a saddlepoint approximation for the distribution function under the discrete case when $\bar{x} > E(x)$, i.e., observed value of the average is greater than the expected. By applying Daniel's approximation we got a saddlepoint approximation for the distribution function for a discrete random variable for $\bar{x} < E(x)$, i.e., observed is less than the expected, and also combined these two results in a more general formulation for the tail (right or left) probability approximation.

To use this saddlepoint approximation result for the distribution function of any standardized U-statistics, we had to approximate the cumulant generating function of both one sample (degree K) and two sample (degree K_1 and K_2) U-statistics in a general framework by using Hoeffding's decomposition of the U-statistics into several uncorrelated U-statistics.

In approximating the cumulant generating function, we tried to keep the order minimal in view of to approximate the tail probability by the saddlepoint methodology (as suggested by I. M. Skovgaard, 1987). This approach enabled us to propose an alternative to the method of approximation for the tail probability of a

standardized U-statistic. The proposed technique has been applied under null and various shift alternatives for normal, logistic and Cauchy distributions for examples in one and two sample settings.

Comparisons through our numerical example were also carried out in the one sample (degree 2) case between the method of Edgeworth expansion as suggested by P. J. Bickel, F. Gotze and W. R. Van Zwet (1986) and the saddlepoint approximation proposed above for the tail probability. Our examples show that tail probability approximation did better than the Edgeworth expansion.

The density of general statistics was also estimated using the saddlepoint approximation by George S. Easton and Elvizio Ronchetti (1986). They approximated the first four cumulants individually and then used these to approximate the cumulant generating function. Using the G. S. Easton and E. Ronchetti's approach, we tried to approximate the cumulant generating function by first four exact cumulants and used it to approximate the tail probability. We then compared the result's of their tail probability approximation with our saddlepoint approximation using an approximated cumulant generating function by applying Hoeffding's decomposition of U-statistics as derived in this dissertation. Our numerical examples show that our method, again, consistently did better over the whole range of critical regions for both one and two sample problems under null and various (normal, logistic and Cauchy) alternatives.

In our examples saddlepoint approximation is closer to exact tail probabilities compared to normal (large sample) approximation. Improvement in the

approximation for these examples is fairly substantial by using saddlepoint approximation. Computing time is negligible since we are using a formula, and obviously, it is much less subjective compared to simulation. Thus, reproduction of results is not a problem. As a result we try to show in this dissertation that it will provide an alternative to normal approximation.

CHAPTER 2
SADDLEPOINT APPROXIMATIONS

Saddlepoint approximations were introduced by H. E. Daniels (1954) in *Annals of Mathematical Statistics.* Based on the theory of complex functions, Daniels (1954) addressed the problem of existence and uniqueness of a saddlepoint. Let X be a random variable with probability function, f(x). Then its moment generating function is defined by

$$M(T) = \int_{-\infty}^{\infty} e^{Tx} f(x)\, d(x)$$

By the inversion formula for characteristic functions we have

$$f(x^*) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M(it)\, e^{-itx^*}\, dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{K(it)-itx^*}\, dt$$

$$= \frac{1}{2i\pi} \int_{-i\infty}^{i\infty} e^{K(T)-Tx^*}\, dT\,. \qquad (2.1)$$

where K(T) is the cumulant generating function of X. Since the density function is expressed in terms of K(T) - Tx^* in the above inversion formula, we need to study the exponent of the integrand carefully. Observe that the exponent is minimized by

T that satisfies the following equation,

$$K'(T) = x^* \quad . \tag{2.2}$$

Existence of real roots for (2.2) has been established under mild and quite general conditions by H. E. Daniels (1954). Daniels also addressed the uniqueness of the saddlepoint without proof. A detailed proof is given in lemma 2.1 and the paragraph following this lemma.

Lemma 2.1: Convexity of cumulant generating function.

Proof: We use Holder's inequality to prove this result, i.e.,

$$E\,|XY| \leq (E\,|X|^p)^{\frac{1}{p}} (E\,|Y|^q)^{\frac{1}{q}}$$

where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

For $\alpha \in (0,1)$ we have

$$M_X(\alpha t_1 + (1-\alpha)t_2)$$

$$= E\, e^{(\alpha t_1 + (1-\alpha)t_2)x}$$

$$= E\left[(e^{t_1 x})^{\alpha} (e^{t_2 x})^{1-\alpha} \right]$$

choosing $p = \frac{1}{\alpha}$ and $q = \frac{1}{1-\alpha}$ and applying Holder's inequality, we have

$$\leq [E\, e^{t_1 x}]^{\alpha} \cdot [E\, e^{t_2 x}]^{1-\alpha}$$

$$= [M(t_1)]^{\alpha} \cdot [M(t_2)]^{1-\alpha}$$

and by taking natural logarithm on both sides we get

$$K(\alpha t_1 + (1-\alpha)t_2) \leq \alpha K(t_1) + (1-\alpha) \cdot K(t_2)$$

and hence this completes the proof of the lemma.

The convexity of cumulant generating function gives us the uniqueness of the minimum of K(T) - Tx* attained at a point T_0 along the real line. Also note, K(T) - Tx* is an analytic function of T(T=x + iy) and thus by Cauchy-Riemann theorem for analytic functions if

$$K(T) - Tx^* = U(x^*, y) + iV(x^*, y)$$

then,

$$\left\{ \begin{array}{l} \dfrac{\partial U}{\partial x} = \dfrac{\partial V}{\partial y} \\ \dfrac{\partial U}{\partial y} = -\dfrac{\partial V}{\partial x} \end{array} \right\} \Rightarrow \frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 V}{\partial x \partial y} = -\frac{\partial^2 U}{\partial y^2} . \qquad (2.3)$$

Also,

$$| e^{K(T) - Tx} | = | e^{U + iV} | = e^U | e^{iV} | = e^U.$$

Along the real line since K(T) - Tx* is real, i.e., V = 0, so that we have

$$\left(\frac{\partial U}{\partial x} \right)_{x = T_0} = 0 \Rightarrow K'(T_0) - x^* = 0$$

and if

$$\left(\frac{\partial^2 U}{\partial x^2} \right)_{x=T_0} = K''(T_0) > 0$$

then, a minimum is attained at T_0 along the real line. Since F(T) = K(T) - Tx is an analytic function, the first derivative is zero along any direction at T_0 and as a result

$$\left(\frac{\partial F}{\partial y} \right)_{\substack{x=T_0 \\ y=0}} = 0 \Rightarrow \left(\frac{\partial U}{\partial y} + \frac{i\partial V}{\partial y} \right)_{\substack{x=T_0 \\ y=0}} = 0$$

$$\Rightarrow \frac{\partial V}{\partial y} = 0 \quad \textit{and} \quad \frac{\partial U}{\partial y} = 0$$

and also from eq. (2.3) we have,

$$\left(\frac{\partial^2 U}{\partial y^2} \right)_{\substack{x=T_0 \\ y=0}} < 0.$$

Therefore, a maximum for U is attained along the imaginary axis and hence T_0, the solution of equation (2.2), is a unique saddlepoint for U.

Let the random variable X take values on a lattice with span = 1 and with probability function $P(x = r) = p_r$. Thus, if M(T) is the moment generating function and k(T) is the cumulant generating function, then we have

$$M(T) = \exp k(T) = \Sigma_r \, p_r \exp (Tr).$$

The probability $P(S = s) = p_n(s)$ for the sum $S = \Sigma X_j$ is given by

$$p_n(s) = \frac{1}{2\pi i} \int_{-i\pi}^{i\pi} e^{n k(T) - Ts} \, d T.$$

Thus the tail probability when s > n E(x) is

$$Q_{n,s} = \sum_{m=s}^{\infty} p_n(m) = \frac{1}{2\pi i} \int_{c-i\pi}^{c+i\pi} e^{n k(T) - Ts} \, \frac{dT}{1 - e^{-T}} \tag{2.3}$$

where c > 0.

Proceeding from eq. (2.3) H. E. Daniels (1987) comes up with a saddlepoint approximation based on Lugannani and Rice's (1980) formula to the tail probability when

$$\bar{x} > E(X) \quad i.e. \quad s > n\,E(X),$$

i.e., the saddlepoint, $\hat{T}$ is positive.

The tail probability is then given by

$$P[S \geq s] = 1 - \Phi(\hat{\xi}) + \phi(\hat{\xi}) \left[\frac{1}{\hat{z}} - \frac{1}{\hat{\xi}} + O(n^{-3/2}) \right] \tag{2.4}$$

where

$$\hat{\xi} = [\, 2n(\hat{T} \cdot \bar{x} - \hat{K}) \,]^{1/2}$$

$$\hat{z} = [\, 1 - \exp(-\hat{T}) \,](n \hat{K}'')^{1/2}$$

where $\hat{K}$ is the cumulant generating function evaluated at $\hat{T}$.

Let Y = -X

then if

$$S = \sum X_j \ , \ S' = \sum Y_j. \qquad note: \ S' = -S.$$

If K(T) is the moment generating function of Y, then K(-T) is the moment generating function of X, i.e.,

$$K_Y(T) = K_X(-T) \ \rightarrow \ K'_Y(T) = -K'_X(-T)$$

$$\rightarrow \ K''_Y(T) = K''_X(-T).$$

Let $\hat{T}_Y \ (> 0)$ be the solution of $K'_Y(T_Y) = s'$ and

$\hat{T}_X \ (< 0)$ be the solution of $K'_X(T_X) = s$.

$$\therefore \quad K'_Y(T) = s' \ \rightarrow \ -K'_X(-T) = -s$$

$$\rightarrow \ K'_X(-T) = s$$

$$\rightarrow \ \hat{T}_Y = -\hat{T}_X$$

and hence we have,

$$\hat{K}_Y = K_Y(\hat{T}_Y) = K_X(-\hat{T}_Y) = K_X(\hat{T}_X) = \hat{K}_X$$

$$\hat{K}'_Y = K'_Y(\hat{T}_Y) = -K'_X(-\hat{T}_Y) = -K'_X(\hat{T}_X) = -\hat{K}'_X \qquad (2.5)$$

$$\hat{K}''_Y = K''_Y(\hat{T}_Y) = K''_X(-\hat{T}_Y) = K''_X(\hat{T}_X) = \hat{K}'_X.$$

From equations (2.4) and (2.5), we get the saddlepoint approximation based on

Lugannani and Rice's formula (1980) to approximate the tail probability when

$$\bar{X} < E(X), \quad \text{i.e.,} \quad S < n \cdot E(X) .$$

I.e., the saddlepoint, $\hat{T}$ is negative.

It is given by

$$P[S' \geq s'] = P[S \leq s] =$$

$$1 - \Phi(\hat{\xi}) + \phi(\hat{\xi}) \left[\frac{1}{\tilde{z}} - \frac{1}{\hat{\xi}} + O(n^{-3/2}) \right] \quad (2.6)$$

where

$$\hat{\xi} = \left[2n(\hat{T}_X \cdot \bar{X} - \hat{K}) \right]^{1/2}$$

$$\tilde{z} = \left[1 - \exp(\hat{T}) \right] (n \hat{K}'')^{1/2}$$

where $\hat{T}$, the saddlepoint is negative and $\hat{K}$ is evaluated at $\hat{T}$.

Equations 2.4 (right tail probability) and 2.6 (left tail probability) can be combined to give a tail probability approximation. It is given as follows:

$$P[Tail] = 1 - \Phi(\xi') + \phi(\xi') \left[\frac{1}{z'} - \frac{1}{\xi'} + O(n^{-3/2}) \right] \quad (2.7)$$

where

$$\xi' = \left[2n(\hat{T} \cdot \bar{X} - \hat{K}) \right]^{1/2}$$

$$z' = \left[1 - \exp(-|\hat{T}|) \right] \cdot (n \hat{K}'')^{1/2} .$$

Since our results in this dissertation are associated with U-statistics (centered about zero) and we, in general, have worked with standardized U-statistics, we felt the need to have a saddlepoint approximation for a tail (right or left) probability of the standardized U-statistics. Let U_N denote a U-statistic with zero mean and σ_N its standard deviation then

$$\sigma_N^{-1} \, U_N \;=\; U_{std}$$

is the standardized U-statistic. We observed that of the two quantities used in this approximation, ξ was scale invariant while z was not. By using an appropriate scale change transformation on z, we further noted that the tail probability approximation based on Lugannani and Rice's (1980) approximation for n = 1 of a standardized U-statistic is given by:

$$P\,[\,Tail\,] \;\sim\; 1 - \Phi\,(\,\xi\,) + \phi\,(\xi\,)\left[\frac{1}{z} - \frac{1}{\xi}\right] \tag{2.8}$$

where

$$\xi \;=\; \left[\, 2(\,\hat{T}\cdot U_{obs} - \hat{K}\,)\,\right]^{1/2}$$

$$z \;=\; \left[\, 1 - \exp\!\left(-\frac{|\hat{T}|}{\sigma_N}\right)\right]\cdot \sigma_N \cdot (\,\hat{K}''\,)^{1/2}\,.$$

Here, $\hat{T}$ is the saddlepoint which is the solution of the saddlepoint equation

$$K'(\,\hat{T}\,) \;=\; U_{std}\,,$$

$K(T)$ is the C.G.F. of the standardized U-statistic and tail means right (left) tail if $U_{std} > 0$ ($U_{std} < 0$).

In the third and fourth chapter, we will look at the problem of approximating the cumulant generating function of a standardized U-statistic for the general one sample (based on a kernel degree K) and two sample case (based on a kernel of degree K_1 and K_2) by using Hoeffding's decomposition (1961) of U-statistics.

CHAPTER 3
ONE SAMPLE: U-STATISTIC

In this chapter we give a brief overview of existing results together with a formal derivation of an approximation for the cumulant generating function of a one sample U-statistic based on a kernel of degree K. This approximation is based on the well known Hoeffding's decomposition (1961) of a U-statistic.

We start in Section 3.1 by developing the notation to be used in this manuscript. We also discuss some known results using this new notation together with their proofs. In some cases, we have given alternate methods of proving these results.

In Section 3.2 we prove two lemmas which allow us to use the martingale structure of a U-statistic to prove that except for a few terms in the expansion based on the Hoeffding's decomposition, everything else can be considered as a high order term. These lemmas are actually an integral part of this dissertation and have been used often in approximations for cumulant generating function of both one sample degree K and two sample degree K_1 and K_2, U-statistics. In addition to these lemmas, we prove the approximation for cumulative generating function of the one sample U-statistics of degree K.

Finally, we discuss an example (in Section 3.3) to which we apply this approximation for cumulant generating function and use it to approximate the tail

probability by using the saddlepoint methodology and comparing these results.

3.1 Notation and Results

Let $X_1, X_2, ..., X_N$ be independent and identically distributed (i.i.d.) random variables with a common distribution F. Let h be measurable and symmetric in its K arguments, i.e.,

$$h(X_1, ..., X_k) = h(X_{i_1}, ..., X_{i_k}).$$

For $N \geq K$, a U-statistic of degree K is defined as

$$U_N = \sum_{1 \leq i_1 < \ldots < i_k \leq N} \cdots \sum h(X_{i_1}, ..., X_{i_k}). \tag{3.1}$$

Note that in the above definition we do not follow the usual convention of dividing the sum in 3.1 by $\binom{N}{K}$ of its terms. Since our results concern the standardized version of a U-statistic, this does not make any difference.

We assume throughout that

$$E\, h(X_1, ..., X_k) = 0, \qquad E\, h^2(X_1, ..., X_k) < \infty.$$

Let

$$g_1(x_1) = E(h(X_1, ..., X_k) \mid X_1 = x_1)$$

$$g_2(x_1, x_2) = E(h(X_1, ..., X_k) \mid X_1 = x_1, X_2, = x_2)$$

$$\vdots \qquad\qquad \vdots$$

$$g_k(x_1, \ldots x_k) = E(h(X_1, ..., X_k) \mid X_1 = x_1, ..., X_k = x_k)$$

$$\psi_1 (x_1) = E (h (X_1, ..., X_k) \mid X_1 = x_1) = g_1 (x_1)$$

$$\psi_2 (x_1, x_2) = g_2 (x_1, x_2) - \left[\sum_{(2,j)} \psi_j (x_j) \right] : j=1.$$

where the sum is over all the number of distinct combinations of size 1 out of 2 X's.

$$\psi_3 (x_1, x_2, x_3) = g_3 (x_1, x_2, x_3) - \left[\sum_{j=1}^{2} \sum_{(3,j)} \psi_j (x_{i_1}, ..., x_{i_j}) \right]$$

. .

. .

. .

$$\psi_k (x_1, ... x_k) = g_k (x_1, ..., x_k) - \left[\sum_{j=1}^{k-1} \sum_{(k,j)} \psi_j (x_{i_1}, ..., x_{i_j}) \right]$$

Also if

$$\zeta_{N-j,k-j} = \frac{(N-j)!}{(N-k)!} = (N-j)(N-j-1)\ldots(N-k+1) = O (N^{k-j})$$

then we further define

$$U_1 = (N-1)(N-2)\ldots(N-k+1) \sum_{i=1}^{N} \psi_1 (X_1) = \zeta_{N-1,k-1} \sum_{i=1}^{N} \psi_1 (X_1)$$

$$U_2 = (N-2)(N-3)\ldots(N-k+1) \sum_{1 \le i_1 < i_2 \le N} \sum \psi_2 (X_{i_1}, X_{i_2})$$

$$= \zeta_{N-2,k-2} \sum_{1 \le i_1 < i_2 \le N} \psi_2 (X_{i_1}, X_{i_2})$$

$$U_3 = (N-3)\ (N-4) \ldots (N-k+1) \sum_{1 \le i_1 < i_2 < i_3 \le N} \psi_3 (X_{i_1}, X_{i_2}, X_{i_3})$$

$$= \zeta_{N-3,k-3} \sum_{1 \le i_1 < i_2 < i_3 \le N} \psi_3 (X_{i_1}, X_{i_2}, X_{i_3})$$

. .

. .

. .

$$U_k = \sum_{1 \le i_1 < i_2 < \ldots < i_k \le N} \psi_k (X_{i_1}, X_{i_2}, \ldots, X_{i_k})$$

$$= \zeta_{N,0} \sum_{1 \le i_1 < i_2 \ldots < i_k \le N} \psi_k (X_{i_1}, X_{i_2}, \ldots, X_{i_k})$$

where $\zeta_{N,0} = 1$.

Now under these definitions for the general one sample U-statistics we try to address some preexisting results which can be expressed under slightly different forms in the literature. This was done for the sake of preserving the continuity and readability of the manuscript.

Let us start with Hoeffding's decomposition for the one sample U-statistics with a kernel of degree k in the following lemma.

Lemma 3.1:

$$U_N = \sum_{j=1}^{k} \frac{U_j}{(k-j)!} . \tag{3.2}$$

Proof:

$$\psi_k (X_1, ..., X_k) = g_k (X_1, ..., X_k) - \sum_{j=1}^{k-1} \left[\sum_{(k,j)} \psi_j (X_1, ..., X_{i_j}) \right]$$

$$= g_k (X_1, ..., X_k) - \sum_{j=1}^{k-1} \left[\sum_{1 \le i_1 < ... < i_j \le k} \psi_j (X_1, ..., X_{i_j}) \right]$$

$$\Rightarrow \psi_k (X_{i_1}, ..., X_{i_k}) = g_k (X_{i_1}, ..., X_{i_k})$$

$$- \sum_{j=1}^{k-1} \left[\sum_{1 \le c_1 < ... < c_j \le k} \psi_j (X_{i_{c_1}}, ..., X_{i_{c_j}}) \right]$$

$$\Rightarrow g_k (X_{i_1}, ..., X_{i_k})$$

$$= \sum_{j=1}^{k} \left[\sum_{1 \le c_i < ... < c_j \le k} \psi_j (X_{i_{c_1}}, ..., X_{i_{c_j}}) \right]$$

Also, we know that

$$U_N = \sum_{1 \le i_1 < ... < i_k \le N} h (X_{i_1}, ..., X_{i_k})$$

and define

$$S_{jN} = \sum_{1 \le i_1 < \dots < i_j \le N} \psi_j (X_{i_1}, \dots, X_{i_j})$$

Then, we have

$$U_N = \sum_{1 \le i_1 < \dots < i_k \le N} \left[\sum_{j=1}^{k} \sum_{1 \le c_1 < \dots < c_j \le k} \psi_j (X_{i_{c_1}}, \dots, X_{i_{c_j}}) \right]$$

$$= \sum_{j=1}^{k} \frac{\binom{N}{k}\binom{k}{j}}{\binom{N}{j}} S_{jN} .$$

A rigorous proof of the above result is given in Serfling (1980), pages 178-179, lemma A (Hoeffding).

Note that,

$$S_{jN} = \sum_{1 \le i_1 < \dots < i_j \le N} \psi_j (X_{i_1}, \dots, X_{i_j})$$

$$= \frac{(N-j) \dots (N-k+1)}{(N-j) \dots (N-k+1)} \sum_{1 \le i_1 < \dots < i_j \le N} \psi_j (X_{i_1}, \dots, X_{i_j})$$

$$= \frac{U_j}{(N-j) \dots (N-k+1)}$$

$$= \frac{N\,(N-1)\,\ldots\,(N-j+1)}{N\,(N-1)\,\ldots\,(N-k+1)} \cdot U_j$$

$$= \frac{{}^{N}P_j}{{}^{N}P_k}\, U_j\,, \qquad \textit{where} \quad {}^{N}P_j = \frac{N\,!}{(N-j)\,!}$$

and hence,

$$U_N = \sum_{j=1}^{k} \frac{\binom{N}{k}\binom{k}{j}\,{}^{N}P_j}{\binom{N}{j}\,{}^{N}P_k} \cdot U_j$$

$$= \sum_{j=1}^{k} \frac{U_j}{(k-j)\,!}$$

This also appears in A. J. Lee's (1990) book, as a theorem, even though the notation is not the same as used here. Now we use the principle of mathematical induction to prove a lemma, which can be used to show that ψ_j's are pair-wise uncorrelated. The following lemma also appears as a theorem with different notation in Lee's book, where an alternative proof is given.

Lemma 3.2: For any k (a positive integer) expectations of the following type are all equal to zero.

$$E\,[\,\psi(\,X_1,\,X_2,\,\ldots\,,\,X_k\,)\,] \;=\; 0$$

$$E\,[\,\psi(X_1, X_2, \ldots, X_k)\,|\,X_1\,] = 0 \tag{3.3}$$

. .

. .

. .

$$E\,[\,\psi(X_1, X_2, \ldots, X_k)\,|\,X_1, \ldots, X_{k-1}\,] = 0$$

Proof: Check for k=1,

$$E\,[\,\psi_1(X_1)\,] = E\,[\,h(X_1, \ldots, X_k)\,] = 0\,.$$

Now, check for k=2,

$$E\,[\psi_2\,(X_1,X_2)] = E\left[\, g_2\,(X_1,X_2) - \sum_{2,j} \psi_j\,(x_j)\,\right] : j=1$$

$$= E\,[\,h\,(X_1, \ldots, X_k)\,] - E\,\psi_1\,(X_1) - E\,\psi_1\,(X_2)$$

$$= 0. \tag{3.4}$$

$$E\,[\,\psi_2\,(X_1, X_2)\,|X_1\,]$$

$$= g_1\,(X_1) - E\,[\,\psi_1\,(X_2)\,|X_1\,) - E\,[\,\psi_1\,(X_1)\,|\,X_1)\,]$$

$$= g_1(X_1) - g_1(X_1) = 0\,. \tag{3.5}$$

Thus, from equations 3.4 and 3.5 we can say that (3.3) holds for k=2.

Assume that (3.3) is true for all $k \leq n$., i.e.,

$$E\,[\,\psi(X_1, X_2, \ldots, X_k)\,] = 0$$

$$E\,[\,\psi(X_1, X_2, \ldots, X_k)\,|\,X_1\,] = 0 \tag{3.6}$$

. .

. .

. .

$$E\,[\,\psi(\,X_1,\,X_2,\,\ldots,\,X_k\,)\mid X_1,\,\ldots,\,X_{k-1}\,]\;=\;0$$

Now, let us check for k=n+1.

$$E\,[\,\psi_{n+1}\,(\,X_1,\,\ldots,\,X_{n+1})\,]\,=\,0.$$

$$E\,[\,\psi_{n+1}\,(X_1,\,\ldots,\,X_{n+1}\,)\mid X_{n+1}\,]$$

$$=\,g_1\,(X_{n+1})\,-\sum_{j=1}^{n}\;\sum_{n+1,j}E\,[\,\psi_j\,(\,X_{i_1},\,\ldots,\,X_{i_j}\,)\mid X_{n+1}\,]$$

$$=\,g_1\,(X_{n+1})\,-\,\{\,E\,[\,\psi_1\,(\,X_{n+1}\,)\mid X_{n+1}\,]\,+\,0\,\}$$

$$\Uparrow$$

from the assumption (3.6) and the fact that X_i's are all i.i.d.

$$=\,g_1\,(X_{n+1})\,-\,g_1\,(X_{n+1})\;=\;0\,.$$

$$E\,[\,\psi_{n+1}\,(X_1,\,\ldots,\,X_{n+1}\,)\mid X_{n+1}\,,\,X_n,\,\ldots,\,X_i\,]\qquad;\qquad 1\,\prec\,i\,\prec\,n{+}1.$$

$$=\,g_{(n+2-i)}(\,X_i,\,X_{i+1}\,,\,\ldots,\,X_{n+1}\,)$$

$$-\sum_{j=1}^{n}\;\sum_{(\,n+1,\,j\,)}^{1}E\,[\,\psi_j\,(\,X_{i_1}\,\ldots\,X_{i_j}\,)\mid X_{n+1},\,\ldots,\,X_i\,]$$

$$=g_{(n+2-i)}\,(X_i,X_{i+1},\,\ldots,\,X_{n+1}\,)$$

$$- \; E\,[\,\psi_{n+2-i}\,(X_{n+1},\,\ldots,\,X_i)\;|\;X_{n+1},\,\ldots,\,X_i\,]$$

$$+ \qquad 0 \qquad = \qquad 0\,.$$

$\Uparrow$

from (3.6), all the previous terms being shown as equal to 0 and the fact that X_i's are all i.i.d.

Hence (3.3) holds for all k (positive integers).

As a consequence of the previous lemma, it is obvious that

$$\psi_i\,(\,X_{j_i}\,,\,\ldots,\,X_{j_i}\,) \quad \text{and} \quad \psi_{i'}\,(\,X_{j_{i'}}\,,\,\ldots,\,X_{j_{i'}}\,) \;\; \textit{for} \;\; 1 \le i \ne i' \le k, \tag{3.7}$$

are pairwise uncorrelated and let

$$\lambda_{N-j,k-j} \;=\; \frac{\zeta_{N-j,k-j}}{(k-j)!} \;=\; \binom{N-j}{k-j},$$

then we have

$$\sigma_N^2 = \mathit{Var}\;(U_N) = \sum_{j=1}^{k} \frac{\mathit{Var}\;(\,U_j\,)}{[\,(\,k-j\,)\,!\,]^2}$$

$$= \frac{[\,(N-1)\,(N-2)\,\ldots\,(N-k+1)\,]^2}{[\,(\,k-1\,)\,!\,]^2}\; N\,E\,g_1^2(\,X_1\,)$$

$$+ \frac{[\,(N-2)\,(N-3)\,\ldots\,(N-k+1)\,]^2}{[\,(\,k-2)\,!\,]^2} \binom{N}{2}\, E\,\psi_2^2(\,X_1,\,X_2\,)$$

$$+ \quad \ldots .$$

$$+ \quad [\, N-(k-1)\,]^2 \cdot \binom{N}{k-1} E\, \psi_{k-1}^2 \,(\, X_1,\, X_2, \ldots,\, X_{k-1}\,)$$

$$+ \quad \binom{N}{k} E\, \psi_k^2 \,(\, X_1,\, X_2,\, \ldots,\, X_k\,)$$

$$\sigma_N^2 = \lambda_{N-1,k-1}^2 \cdot N \cdot E\, g_1^2 (\, X_1\,)$$

$$+ \lambda_{N-2,K-2}^2 \cdot \binom{N}{2} \cdot E\, \psi_2^2 (\, X_1,\, X_2\,)$$

$$+ \ldots + \lambda_{N-k+1,1}^2 \binom{N}{k-1} E\, \psi_{k-1}^2 (\, X_1,\, X_2,\, \ldots,\, X_{k-1}\,)$$

$$+ \lambda_{N-k,0}^2 \binom{N}{k} E\, \psi_k^2 \,(\, X_1,\, X_2,\, \ldots,\, X_k\,) \qquad (3.8)$$

[$\because$ *Xi's are all i.i.d.*].

If it is assumed that $\sigma_{g_1}^2 = E\, g^2\,(X_1) > 0$, then first term dominates the right-hand side of equation 3.8 and $\sigma_N^{-1}\, U_N$ is asymptotically normal as $N \to \infty$ (c.f. Hoeffding, 1948, where U-statistics were introduced). Also note that

$$\sigma_N^2 = O\,(\, N^{2k-1}\,)\,,$$

i.e., $$\sigma_N = O\left(N^{k-1/2}\right) \tag{3.9}$$

[$\because$ X_i's are all i.i.d.]

3.2 U-Statistic: Degree K

We now prove two lemmas which are needed to derive the approximation for the cumulant generating function of both one and two sample U-statistics. We use Hoeffding's decomposition in both these derivations.

Lemma 3.3 :

Assume $E \mid h(X_1, \ldots, X_k) \mid^{2+\delta} < \infty$ for some $\delta \geq 0$,

and further define $S_{j,N} = \sum_{1 \leq i_1 < \ldots < i_j \leq N} \cdots \sum \psi_j(X_{i_1}, X_{i_2}, \ldots, X_{i_j})$

then

$$E\left(\mid S_{j,N} \mid^{2+\delta}\right) \leq O\left(N^{j(1+\delta/2)}\right) \tag{3.10}$$

Proof: Let us consider

$$S_{2,N} = \sum_{1 \leq i_1 < i_2 \leq N} \cdot \sum \psi_2(X_{i_1}, X_{i_2})$$

$$= \sum_{i_2=2}^{N} \sum_{i_1=1}^{i_2-1} \psi_2(X_{i_1}, X_{i_2})$$

$$= \sum_{i_2=2}^{N} S_{1,i_2-1} \quad (say)$$

Where $S_{1,0} = 0$ and

$$E\ (\ S_{1,i_2-1} \mid S_{1,0}\ ,\ \ldots\ ,\ S_{1,i_2-2}\ ) = 0 \qquad \forall \qquad i_2 = 2,3,4\ldots,\ N.$$

By lemma 3.2.

Thus if $S_{2,1} = 0$, then

$$S_{2,k_2} = \sum_{i_2=2}^{k_2} S_{1,i_2-1} \quad ; \quad k_2 = 2,\ldots,N.$$

has a martingale structure and S_{1,i_2-1} are its summands. Using a result of Dharmadhikari, Fabian and Jogdeo (1968) on bounds on the moments of martingales for $k_2 = N$ and $C_2 =$ constant, we get

$$E\ (\ |\ S_{2,N}\ |^{2+\delta}\ ) \leq C_2 \cdot (N-1)^{1+\delta/2} \cdot \max_{2 \leq i_2 \leq N} E\ (\ |S_{1,i_2-1}\ |^{2+\delta}\ ) \qquad (3.11)$$

but for fixed $i_2 \geq 2$

$$S_{1,k} = \sum_{i_1=1}^{k} \psi_2\ (\ X_{i_1}\ ,\ X_{i_2}\ ) \quad ; \quad k = 1\ ,2\ ,\ \ldots.,\ i_2 - 1.$$

is also a martingale and by applying Dharmadhikari, Fabian and Jogdeo's (1968) result, again we have

$$E\,(\,|\,S_{1,i_2-1}\,|^{2+\delta}\,) \le C_1 \cdot (i_2-1)^{1+\delta/2} \cdot \max_{1 \le i_1 \le i_2-1} E\,|\,\psi_2(\,X_{i_1}\,,\,X_{i_2}\,)\,|^{2+\delta}$$

$$= C_1 \cdot (i_2-1)^{1+\delta/2} \cdot E\,|\,\psi(\,X_1\,,\,X_2\,)\,|^{2+\delta}$$

$$= C_1 \cdot (i_2-1)^{1+\delta/2} \cdot E\,|\,g_2(\,X_1,\,X_2\,) -$$

$$- g_1(\,X_1\,) - g_1(\,X_2\,)\,|^{2+\delta}$$

$$\le C_1' \cdot (i_2-1)^{1+\delta/2} \cdot \Big[\, E\,|\,g_2(\,X_1,\,X_2\,)\,|^{2+\delta} +$$

$$+ E\,|\,g_1(\,X_1\,)\,|^{2+\delta} + E\,|\,g_1(\,X_2\,)\,|^{2+\delta}\,\Big]$$

from C_r inequality (C. R. Rao, 1973).

$$\le C_1'' \cdot (i_2-1)^{1+\delta/2} \cdot E\,|\,h(\,X_1,\,\ldots,\,X_k\,)\,|^{2+\delta} \tag{3.12}$$

from theorem 9.1.4 (K. L. Chung, 1968),

and thus from equations 3.11, 3.12 and $E\,|\,h(\,X_1,\,X_2,\,\ldots.,\,X_k\,)\,|^{2+\delta} < \infty$ we get,

$$E\,|\,S_{2,N}\,|^{2+\delta} \le O(\,N^{2+\delta}\,) \tag{3.13}$$

Again consider,

$$S_{3,N} = \sum_{1 \le i_1 < i_2 < i_3 \le N} \cdots \sum \psi_3(\,X_{i_1}\,,\,X_{i_2},\,X_{i_3}\,)$$

$$= \sum_{i_3=3}^{N} \sum_{i_2=2}^{i_3-1} \sum_{i_1=1}^{i_2-1} \psi_3(\,X_{i_1}\,,\,X_{i_2}\,,\,X_{i_3}\,)$$

$$= \sum_{i_3=3}^{N} S_{2,i_3-1}$$

Note that we already know that S_{2,i_3-1} ; $i_3 = 3, 4, \ldots, N.$ has a martingale structure and furthermore,

$$E\ (\ S_{2,i_3-1} \mid S_{2,1}\ ,\ \ldots\ ,\ S_{2,i_3-2}\ )\ =\ 0 \qquad \forall \qquad i_3 = 3,4,\ldots,\ N.$$

by lemma 3.2.

$$\therefore\ S_{3,k_3} = \sum_{i_3=3}^{k_3} S_{2,i_3-1} \quad ; \quad k_3 = 3\ ,\ldots,\ N.$$

also has a martingale structure and $S_{3,1} = S_{3,2} = 0$. Thus by again applying Dharmadhikari, Fabian and Jogdeo's (1968) result for $k_3 = N$ we have,

$$E \mid S_{3,N} \mid^{2+\delta} \ \le\ C_3 \cdot (N-2)^{1+\delta/2} \cdot \max_{3 \le i_3 \le N} E(\ \mid S_{2,i_3-1} \mid^{2+\delta}\ )$$

where C_3 is a constant, but from (3.13) we have,

$$E\ (\ \mid S_{2,i_3-1} \mid^{2+\delta}\ )\ \le\ O(\ \mid i_3 - 1 \mid^{2+\delta}\ )$$

$$\rightarrow \qquad \max_{3 \le i_3 \le N} E\ (\ \mid S_{2,i_3-1} \mid^{2+\delta}\ )\ \le\ O(\ N^{2+\delta}\ )$$

and hence

$$E\ (\ \mid S_{3,N} \mid^{2+\delta}\ )\ \le\ O(\ N^{3(1+\delta/2)}\ ). \tag{3.14}$$

Now let us assume that

$$E\ (\ |\ S_{j-1,N}|^{2+\delta}\ )\ \le\ O(\ N^{(j-1)(1+\delta/2)}\ ) \tag{3.15}$$

holds for some j > 2. Now consider,

$$S_{j,N} = \sum_{1\le i_1<\ldots<i_j\le N} \cdots \sum \psi_j(\ X_{i_1}\ ,\ X_{i_2},\ \ldots.,\ X_{i_j}\ )$$

$$= \sum_{i_j=j}^{N} \sum_{i_{j-1}=j-1}^{i_j-1} \cdots \sum_{i_1=1}^{i_2-1} \psi_j(\ X_{i_1}\ ,\ X_{i_2}\ ,\ \ldots.,\ X_{i_j}\ )$$

$$= \sum_{i_j=j}^{N} S_{j-1,i_j-1}$$

Note also that if $S_{j-1,1} = S_{j-1,2} = \ldots = S_{j-1,j-1} = 0$, then

$$E\ (\ S_{j-1,i_j-1}\ |\ S_{j-1,1}\ ,\ \ldots\ ,\ S_{j-1,i_j-2}\ )\ =\ 0 \qquad \forall \quad i_j = j,\ j+1,\ \ldots\ ,\ N.$$

by lemma 3.2.

Thus

$$S_{j,k_j} = \sum_{i_j=1}^{k_j} S_{j-1,i_j-1} \qquad ; \qquad i_j\ ,\ k_j = j\ ,\ j+1,\ \ldots\ ,\ N.$$

has a martingale structure. Again using Dharmadhikari, Fabian and Jogdeo's (1968) result for $k_j = N$ we have,

$$E\ (\ |\ S_{j,N}\ |^{2+\delta}\ \le\ C_5\ \cdot\ (\ N-j+1\ )^{1+\delta/2}\ \cdot \max_{j\le i_j\le N}\ E\ (\ |\ S_{j-1,i_j-1}\ |^{2+\delta}\ )$$

$$\leq O\left(N^{j(1+\delta/2)} \right) \qquad \text{from (3.15).}$$

Therefore, we know that if equation 3.10 holds for j-1, then it also holds for j and hence by principle of mathematical induction, equation 3.10 holds for any j.

Lemma 3.4 :

Under $E \mid h\left(X_1, \ldots, X_k \right) \mid^{2+\delta} < \infty$ for some $\delta \geq 0$,

then we have

$$E \mid U_{-i} \mid^{2+\delta} = O\left(N^{[k-(i+1)/2](2+\delta)} \right) \tag{3.16}$$

Proof:

$$\left[E \mid U_{-i} \mid^{2+\delta} \right]^{\frac{1}{2+\delta}} = \left[E \mid U_{i+1}/(k-i-1)! + \ldots + U_k \mid^{2+\delta} \right]^{\frac{1}{2+\delta}}$$

$$\leq \left[E \mid U_{i+1}/(k-i-1)! \mid^{2+\delta} \right]^{\frac{1}{2+\delta}} + \ldots + \left[E \mid U_k \mid^{2+\delta} \right]^{\frac{1}{2+\delta}}$$

by using the Minkowski inequality.

Let $h\left(X_1, \ldots, X_k \right)$ be the kernel of degree k for an appropriate U-statistic and let us consider,

$$E \mid U_j \mid^{2+\delta} = E \mid (N-j)\ldots(N-k+1) \sum_{1 \leq i_1 < \ldots < i_j \leq N} \cdots \sum \psi_j \left(X_{i_1}, \ldots, X_{i_j} \right) \mid^{2+\delta}$$

$$= O\left(N^{(2+\delta)(k-j)} \right) \cdot E \mid \sum_{1 \leq i_1 < \ldots < i_j \leq N} \cdots \sum \psi_j \left(X_{i_1}, \ldots, X_{i_j} \right) \mid^{(2+\delta)}$$

$$\leq O\left(N^{(2+\delta)(k-j)}\right) \cdot O\left(N^{\left(\frac{j}{2}\right)(2+\delta)}\right)$$

from lemma 3.3.

$$= O\left(N^{(2+\delta)(k-j/2)}\right) \leq O\left(N^{[k-(i+1)/2](2+\delta)}\right)$$

where j = i+1, i+2,..., k.

Note that our main assumption in the derivations of lemmas 3.3 and 3.4 was

$$E \mid h(X_1, \ldots, X_k) \mid^{2+\delta} < \infty$$

where $\delta \geq 0$.

We will now use the above two lemmas together with the Hoeffding's decomposition to derive the approximation for the cumulant generating function of a general one sample standardized U-statistic based on a kernel of degree K. Here, we assume $\delta \in (0,1]$ in the previous two lemmas. Let us consider

$$\phi_N(t) = E \exp\left(t \frac{U_N}{\sigma_N}\right) = E \exp\left(t \sum_{j=1}^{k} \frac{U_j}{(k-j)!\,\sigma_N}\right)$$

$$= E \exp\left[\frac{t}{\sigma_N}\left(\frac{U_1}{(k-1)!} + \frac{U_2}{(k-2)!} + \ldots + \frac{U_k}{1}\right)\right]$$

$$= E\left\{\exp\left[t \frac{\lambda_{N-1,k-1} \sum_{i=1}^{N} \psi_1(X_i)}{\sigma_N}\right] \exp\left[t \frac{\sum_{j=2}^{k} \frac{U_j}{(k-j)!}}{\sigma_N}\right]\right\}$$

$$= \prod_{i=1}^{N} E \exp\left(t \frac{\lambda_{N-1,k-1}\, \psi_1 (X_1)}{\sigma_N} \right)$$

$$+ E\left[\prod_{i=1}^{N} \exp\left(t \frac{\lambda_{N-1,k-1}\, \psi_1 (X_i)}{\sigma_N} \right) \times \right.$$

$$\left. t \frac{(U_2/(k-2)! + U_3/(k-3)! + U_{-3})}{\sigma_N} \right]$$

$$+ E\left[\prod_{i=1}^{N} \exp\left(t \frac{\lambda_{N-1,k-1}\, \psi_1 (X_i)}{\sigma_N} \right) \cdot \frac{\left(t \frac{(U_2/(k-2)! + U_{-2})}{\sigma_N} \right)^2}{2} \right.$$

$$+ O\left(E \left| t \frac{U_{-1}}{\sigma_N} \right|^{2+\delta} \right) \quad \text{for any } \delta \in (0,1]. \tag{3.17}$$

where $U_{-i} = \sum_{j=i+1}^{k} \frac{U_j}{(k-j)!}$

Note that from lemma 3.4 we have

$$E\, |U_{-i}|^{2+\delta} = O\left((N^{k-(i+1)/2})^{2+\delta} \right).$$

for $\delta \geq 0$.

and also observe that

$$E\left[\prod_{i=1}^{N} \exp\left(t\,\frac{\lambda_{N-1,k-1}\,\psi_1(X_i)}{\sigma_N}\right)\cdot t\,\frac{U_{-3}}{\sigma_N}\right] = O(N^{-\frac{3}{2}}). \tag{3.18}$$

and the interaction term within the square term in equation 3.17 is also of the order

$$O(N^{-3/2}) \tag{3.19}$$

Let

$$\gamma_N(t) = E\exp\left(t\,\frac{\delta_{N-1,k-1}\,\psi_1(X_1)}{\sigma_N}\right)$$

then from equations 3.18, 3.19 and lemma 3.4 we have,

$$\phi_N(t) = \gamma_N^N(t)$$

$$+\,\gamma_N^{N-2}(t)\cdot t\cdot\frac{\lambda_{N-2,k-2}}{\sigma_N}\binom{N}{2}\left(t\,\frac{\lambda_{N-1,k-1}}{\sigma_N}\right)^2 \times$$

$$E\,\psi_1(X_1)\,\psi_1(X_2)\,\psi_2(X_1,X_2)$$

$$+\,\gamma_N^{N-3}(t)\cdot t\cdot\frac{\lambda_{N-3,k-3}}{\sigma_N}\binom{N}{3}\left(t\,\frac{\lambda_{N-1,k-1}}{\sigma_N}\right)^3$$

$$\cdot\,E\,\psi_1(X_1)\,\psi_1(X_2)\,\psi_1(X_3)\,\psi_3(X_1,X_2,X_3)$$

$$+ \gamma_N^{N-2}(t) \cdot t \frac{\lambda_{N-2,k-2}}{\sigma_N} \binom{N}{2} \left(t \frac{\lambda_{N-1,k-1}}{\sigma_N} \right)^3 E\, \psi_1^2(X_1)\, \psi_1(X_2)\, \psi_2(X_1,X_2)$$

$$+ \gamma_N^{N-2}(t) \cdot t^2 \cdot \frac{\sigma_N^{-2}}{2} \cdot \lambda_{N-2,k-2}^2 \binom{N}{2} E\, \psi_2^2(X_1,X_2)$$

$$+ 6 \cdot \gamma_N^{N-3} \cdot t^2 \cdot \frac{\lambda_{N-2,k-2}^2}{2 \cdot \sigma_N^2} \binom{N}{3} \left(t \frac{\lambda_{N-1,k-1}}{\sigma_N} \right)^2$$

$$\cdot E\, \psi_1(X_1)\, \psi_1(X_2)\, \psi_2(X_1,X_3)\, \psi_2(X_2,X_3)$$

$$+ 6 \cdot \gamma_N^{N-4}(t) \cdot t^2 \cdot \frac{\lambda_{N-2,k-2}^2}{2 \cdot \sigma_N^2} \binom{N}{4} \left(t \frac{\lambda_{N-1,k-1}}{\sigma_N} \right)^4$$

$$\cdot \left[E\, \psi_1(X_1)\, \psi_1(X_2)\, \psi_2(X_1,X_2) \right]^2$$

$$+ 0\left(|N^{-\frac{1}{2}} t|^{2+\delta} \right) \qquad (3.20)$$

This expansion of the cumulant generating function for the one sample U-statistics with a kernel of degree k, reduces to P. J. Bickel, F. Gotze and W. R. Van Zwet's (1986) expansion of the cumulant generating function for one sample degree two case. In fact, there is only one additional term in this expansion. This additional term of order $O(N^{-1})$ is important for U-statistics with kernel of degree

≥ 3, i.e., if order is equal to two, this term dissappears.

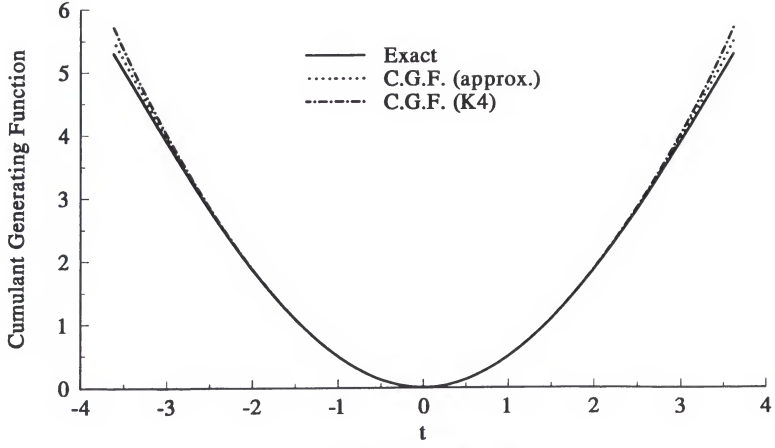
3.3. Example

We used the Wilcoxon Signed Rank test statistic as an example to test the accuracy of the cumulant generating function approximation as defined on the previous page. Note that the signed rank test statistic is actually a linear combination of two U-statistics. Since it is possible to get hold of the exact cumulant generating function of the signed rank statistic, it was easy to carry out this comparison, i.e., see Figure 3.1. Cumulant generating function based on the first four (exact) cumulants has also been included in the comparison as we wanted to investigate the behavior of resulting saddlepoint approximation, Easton and Ronchetti (1986), based on it also. We were also interested if eq. (3.20) could be successfully used to approximate the tail probability by the saddlepoint approximation as given in equation 2.8 (Chapter 2) and at the same time also provide a viable alternative to large sample (normal) approximation.

To use Hoeffding's decomposition we consider the Wilcoxon Signed Rank test statistic. Let $X_1, \ldots, X_N$ be i.i.d. random variables with a common continuous distribution function F. Let

$$R_i^+ = \sum_{j=1}^{N} 1_{[\,|x_j| \leq |x_i|\,]} \tag{3.21}$$

be the signed rank of X_i and also let W_N^+ denote signed rank test statistic for



Wilcoxon Signed Rank Test
Using Exact, C.G.F. (approx.) and C.G.F. (K4)
Comparison of Cumulant Generating Function

Figure 3.1

testing the hypothesis that the distribution of X_1 is symmetric about zero, thus

$$W_N^+ = \sum_{i=1}^{N} 1_{[X_i \geq 0]} \cdot R_i^+ \tag{3.22}$$

Note that $\sum_{i=1}^{j} \psi (X_{(i)} + X_{(j)})$ is just the signed rank of $X_{(j)}$ where $X_{(i)}$ is the i^{th} order

statistic in the sample and

$$\psi(X) = \begin{cases} 1 & if \ X > 0 \\ 0 & otherwise. \end{cases} \tag{3.23}$$

Let us consider

$$W_N^+ = \sum_{1 \leq i \leq j \leq N} \cdots \sum \psi(X_{(i)} + X_{(j)}) = \sum_{1 \leq i \leq j \leq N} \cdots \sum \psi (X_i + X_j)$$

$$= \sum_{j=1}^{N} \psi (2 X_j) + \sum_{1 \leq i < j \leq N} \cdots \sum \psi (X_i + X_j)$$

$$= N \cdot U_3(X_1, \ldots , X_N) + \binom{N}{2} \cdot U_4(X_1, \ldots , X_N)$$

where

$$U_3 (X_1, \ldots, X_N) = \frac{1}{N} \sum_{i=1}^{N} \psi (X_i)$$

and $$U_4 (X_1, \ldots, X_N) = \frac{1}{\binom{N}{2}} \sum_{1 \le i < j \le N} \cdots \sum \psi (X_i + X_j)$$

$$\Rightarrow \quad E (W_N^+) = N P (X_j > 0) + \frac{N (N-1)}{2} [P (| X_i | < | X_j | , X_j > 0)$$

$$+ P (| X_j | < | X_i | , X_i > 0)]$$

$$= N (1 - F(0)) + N (N-1) \int_0^\infty P (-x < X_i < x) \, dF(X)$$

$$= N (1 - F(0)) + N (N-1) \int_0^\infty (F (x) - F (-x)) \, dF(x)$$

Now let us consider,

$$U_N = W_N^+ - E W_N^+$$

$$= \sum_{j=1}^{N} \psi (X_j) + \sum_{1 \le i < j \le N} \cdots \sum \psi (X_i + X_j)$$

$$- N (1 - F (0)) - N (N-1) \int_0^\infty (F (x) - F (-x)) \, dF(x) . \quad (3.24)$$

Then U_N is clearly a U-statistic, and it is well known that it can be written (by using

Hoeffding's decomposition) as follows:

$$U_N = (N-1) \sum_{i=1}^{N} g_1 (X_i) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \psi_2 (X_i, X_j) , \tag{3.25}$$

where

$$g_1 (x) = (1 - F (-x)) - \int_{-\infty}^{\infty} (1 - F (-x) \, dF (x) +$$

$$\frac{1}{N-1} [1_{[0,\infty)}(x) - 1 + F (0)]$$

and

$$\psi_2 (x, y) = 1_{[0,\infty]}(x+y) - (1 - F (-x)) - (1 - F (-y)) +$$

$$\int_{-\infty}^{\infty} (1 - F (-x)) \, dF(x) .$$

Under this setup, B. J. Bickle, F. Gotze and W. R. van Zwet (1986) defined

$$\gamma_N (t) = E \exp [t \, \sigma_N^{-1} (N-1) \, g (X_1)],$$

which is the moment generating function of $\sigma_N^{-1} (N-1) \, g (X_1)$

$$\hat{U}_N = (N-1) \sum_{i=1}^{N} g (X_i) \quad \text{and} \quad \Delta_N = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \psi (X_i , X_j) .$$

where X_i's are all independent and identically distributed random variables.

Then they derived the following approximation to the moment generating function of Wilcoxon Signed Rank test statistic.

$$\Phi_N(t) = \gamma_N^N(t)$$

$$+ \gamma_N^{N-2}(t)\left[\frac{1}{2}\cdot t^3\,\sigma_N^{-3}\,N\,E\,g(X_1)\,g(X_2)\,\psi(X_1,X_2)\right.$$

$$+ t^4\,\sigma_N^{-4}\,N^5\,E\,g^2(X_1)\,g(X_2)\,\psi(X_1,X_2)$$

$$\left. + \frac{1}{4}\cdot t^2\,\sigma_N^{-2}\,N^2\,E\,\psi^2(X_1,X_2)\right]$$

$$+ \frac{1}{2}\cdot\gamma_N^{N-3}(t)\,t^4\,\sigma_N^{-4}\,N^5\times$$

$$E\,g(X_1)\,g(X_2)\,\psi(X_1,X_3)\,\psi(X_2,X_3)$$

$$+ \frac{1}{8}\gamma_N^{N-4}(t)\,t^6\,\sigma_N^{-6}\,N^8\left[E\,g(X_1)\,g(X_2)\,\psi(X_1,X_2)\right]^2$$

$$+ O\left(|\gamma_N(t)|^{N-4}\,P(|t|)\,N^{-1-3\epsilon} + |N^{-\frac{1}{2}}t|^{2+\delta}\right) \qquad (3.26)$$

where P is a polynomial.

Note that this result is a special case of our approximation to the moment generating function of one sample U-statistics based on a kernel of degree k (i.e.,

where k = 2) given earlier in this chapter (equation 3.20).

We used the result from the equation 3.20 (for approximating the moment generating function) to further approximate the tail probability by the saddlepoint methodology as given in Chapter 2, equation (2.8). To do this we first had to approximate four expectations under the null hypothesis. Any standard numerical integration technique / package could have been used for this approximation. Since U-statistics are distribution free, these expectations are independent of the underlying distributions as long as the distributions are symmetric about zero. The approximated values of these expectations were:

(i) $E_1 = E\, g(X_1)\, g(X_2)\, \psi(X_1, X_2) = 0$,

(ii) $E_2 = E\, g^2(X_1)\, g(X_2)\, \psi(X_1, X_2) = -0{\cdot}00346951$, (3.27)

(iii) $E_3 = E\, g(X_1)\, g(X_2)\, \psi(X_1, X_3)\, \psi(X_2, X_3) = 0{\cdot}00173457$,

(iv) $E_4 = E\, \psi^2(X_1, X_2) = 0{\cdot}083333$.

Note also that except for these expectations everything else in equation 3.20 (for k = 2) is either a constant or a function of t. Thus, when we use eq. 3.20 to approximate the cumulant generating function (of signed rank test statistic) and henceforth equation 2.2 in Chapter 2 to evaluate the saddlepoint (which we need for saddlepoint approximation), we end up with a function of t, which needs to be minimized with respect to t. The value of t for which this function is minimized is the saddlepoint.

Besides the normal approximation we also computed two other saddlepoint

approximations.

The first approximation was based on the exact cumulant generating function and since the tail probability based on this approximation was exactly the same to that of the null distribution of Wilcoxon Signed Rank test statistic, see Table 3.1, this tail probability approximation was excluded from Figure 3.2 to keep the number of comparisons optimal.

For the second saddlepoint approximation, we used an approach of G. S. Easton and E. Ronchetti (1986) where they proposed to approximate the first four cumulants of the statistic and in turn, used these to further approximate the cumulant generating function. used in the saddlepoint approximation. Here, mainly to optimize their approximation, we used exact cumulants (the first four) to approximate the cumulant generating function used in the saddlepoint approximation. Results of this approximation have both been tabulated in Table 3.1 and plotted in Figure 3.2 (A-C).

From the table it is clear that the saddlepoint approximation works quite well to approximate the tail probability under the null distribution for the Wilcoxon Signed Rank test statistic. In fact, it is very close (up to the fourth decimal place) to the null distribution (exact) if exact cumulant generating function is used, see Table 3.1.

Saddlepoint approximation based on the cumulant generating function approximated by the Hoeffding's decomposition also works quite well. It not only beats the normal approximation over almost the whole range, but it also beats the

Table 3.1. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) between Exact (Null) Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Sp_app	Exact	Sp_ex	Sp_k_4	Normal
35	-1.964530	0.0262973	0.0253	0.0252818	0.0304044	0.0247
40	-1.727840	0.0449721	0.0443	0.0443508	0.0470689	0.0420
45	-1.491150	0.0731930	0.0727	0.0727472	0.0746681	0.0680
50	-1.254460	0.1127630	0.1123	0.1124000	0.1137430	0.1048
55	-1.017770	0.1648830	0.1645	0.1645580	0.1654000	0.1544
60	-0.781079	0.2296940	0.2293	0.2293990	0.2298360	0.2174
65	-0.544388	0.3060150	0.3056	0.3057790	0.3059420	0.2931
70	-0.307698	0.3913040	0.3910	0.3911510	0.3911830	0.3792
75	-0.071007	0.4818850	0.4816	0.4818470	0.4818480	0.4717

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

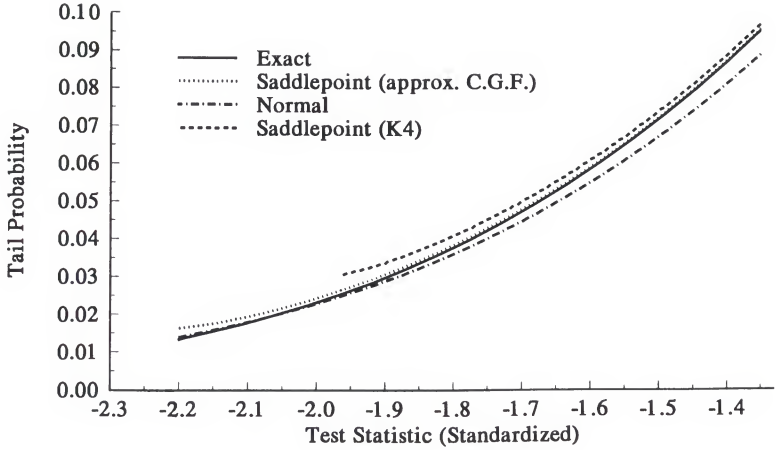
Sp_app = Saddlepoint Approximation Based on Our Approximation

Exact = Null Distribution of Test-Statistic

Sp_ex = Saddlepoint Approximation Based on Exact C.G.F.

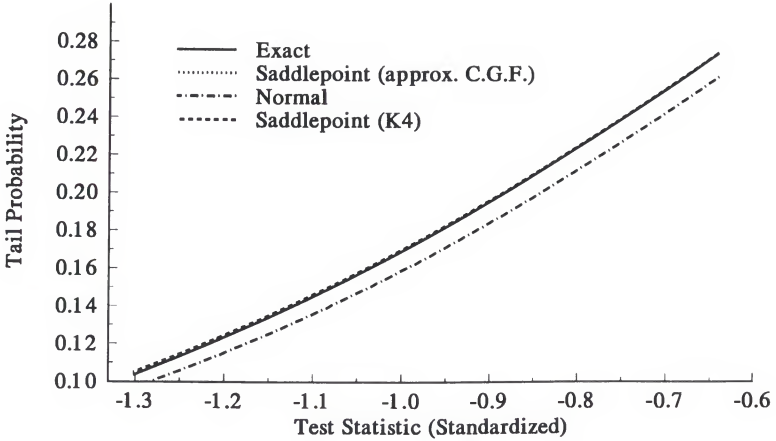
Sp_K_4 = Saddlepoint Approximation Based on Exact Cumulants (K_1-K_4)

Normal = Normal Approximation

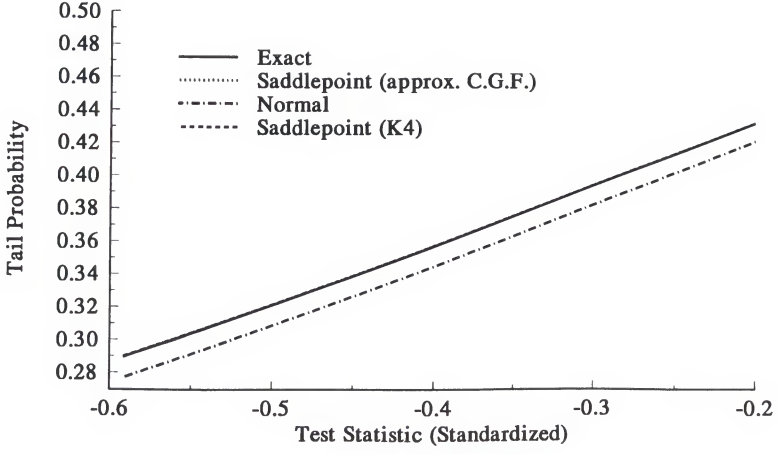


Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution

Figure 3.2 (A)



Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution
Figure 3.2 (B)



Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution
Figure 3.2 (C)

saddlepoint approximation as proposed by Easton and Ronchetti (1986) in JASA, where the cumulant generating function has been approximated by using the first four cumulants (exact).

For an extreme test statistic value, when the tail probability becomes very small, i.e., the standardized test statistic becomes quite large, and as a result the saddlepoint also relatively becomes large and then the saddlepoint approximation (based on the cumulant generating function based on Hoeffding's decomposition) tends to lose its accuracy. The reason for this is that for a large value of t, the cumulant generating function approximated by the decomposition theorem may not be too accurate and hence the saddlepoint itself may be less accurate. If this becomes a problem, we suggest that some of the higher order terms may need to be considered. It may be worthwhile to note that this is not a concern for saddlepoint approximation based on exact cumulant generating function, which beats the normal approximation over the entire range.

CHAPTER 4
TWO SAMPLE: U-STATISTICS

In this chapter, we look at the two sample U-statistics and formally derive the approximation for cumulant generating function (similar to the one sample case, discussed in Chapter 3) based on the Hoeffding's decomposition.

We start by giving a detailed proof for two sample U-statistics based on a kernel of degree (1,1) in Section 4.1. This will not only allow it to be used directly for approximation of cumulant generating function for any two sample U-statistics of degree (1,1) such as the Mann-Whitney-Wilcoxon statistic, two sample proportion problem, etc., but also make the proof for the general case easier to understand.

In Section 4.2 we briefly discuss the derivation of an approximation for the cumulant generating function of a general two sample U-statistic based on a kernel of degree K_1 and K_2. The proof discussed in this section is an extension of the one sample U-statistic degree K derived in the previous chapter. We also apply the lemmas proved in Chapter 3.

Finally, we discuss an example (in section 4.3) to which we apply this approximation for cumulant generating function and use it to approximate the tail probability by using the saddlepoint methodology and compare the results.

4.1 Two Sample: Degree (1,1)

Let $X_1, X_2, ..., X_m$ be independently and identically distributed (i.i.d.) random variables with distribution function F(X) and let $Y_1, Y_2, ..., Y_n$ be i.i.d. random variables with distribution function G(Y). If h is a measurable function of two arguments

$$h (X_i, Y_j)$$

which is trivially symmetric within its X's and Y's, then a two sample U-statistics ($U_{m,n}$) based on a kernel h, of degree (1,1) is defined as follows:

$$U_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} h (X_i, Y_j) \qquad (4.1)$$

Note that again we consider a standardized version of two sample U-statistics of degree (1,1), where we assume without loss of generality that

$$E\, h (X_1, Y_1) = 0 \qquad \textit{and} \qquad E\, h^2 (X_1, Y_1) < \infty .$$

Notation & Assumption:

Let us define,

$$\hat{U}_n = m \sum_{j=1}^{n} g_2 (Y_j) = m \sum_{j=1}^{n} E \left(h (X_i, Y_j) \middle| Y_j \right)$$

$$\hat{U}_m = n \sum_{i=1}^{m} g_1 (X_i) = n \sum_{i=1}^{m} E \left(h (X_i, Y_j) \mid X_i \right)$$

$$\Delta_{mn} = \sum_{i=1}^{m} \sum_{j=1}^{n} \psi (X_i, Y_j) \qquad \textit{where}$$

$$\psi (x, y) = h (x, y) - g_1 (x) - g_2 (y)$$

Such that

$$U_{m,n} = \hat{U}_n + \hat{U}_m + \Delta_{mn} .$$

Since

$$E (\psi(X_1, Y_1) \mid X_1) = 0 \quad \textit{and} \quad E (\psi(X_1, Y_1) \mid Y_1) = 0 ,$$

the random variables $g_1(X_i)$, $g_2(Y_j)$ and $\psi(X_i, Y_j)$, $1 \le i \le m$, $1 \le j \le n$, are pairwise uncorrelated and hence

$$\sigma^2_{m,n} = Var (U_{m,n}) = \sigma^2 (U_{m,n})$$

$$= \sigma^2 (\hat{U}_n) + \sigma^2 (\hat{U}_m) + \sigma^2 (\Delta_{mn})$$

$$= m^2 n E g_2^2 (Y_1) + n^2 m E g_1^2 (X_1) + n m E \psi^2 (X_1, Y_1) .$$

If it is assumed that

$$\sigma_{g_1}^2 = E g_1^2 (X_1) > 0 \quad \textit{and} \quad \sigma_{g_2}^2 = E g_2^2 (Y_1) > 0 ,$$

then $\sigma^2 (\hat{U}_m)$ and $\sigma^2 (\hat{U}_n)$ dominate the right hand side of (4.) and also $\sigma_N^{-1} U_N$ is asymptotically normal as $n, m \to \infty$ (c.f. Hoeffding, 1948). If we further assume that there exists a $c > 0$, such that $m = c \cdot n$, then we have

$$\sigma^2_{m,n} = O(n^{-3}) \quad i.e. \quad \sigma_{m,n} = O(n^{-\frac{3}{2}}) .$$

and similar to lemma 3.4. it can be shown that

$$E \mid \Delta_{mn} \mid^{2+\delta} = O(n^{2+\delta}) .$$

Let

$$\gamma_m(t) = E \exp\{ t\, \sigma^{-1}_{m,n}\, m\, g_2(Y_1) \}$$

be the moment generating function of $t\, \sigma^{-1}_{m,n}\, m\, g_2(Y_1)$ and

$$\gamma_n(t) = E \exp\{ t\, \sigma^{-1}_{m,n}\, n\, g_1(X_1) \}$$

be the moment generating function of $t\, \sigma^{-1}_{m,n}\, n\, g_1(X_1)$.

Now let us consider,

$$\phi_{m,n}(t) = E \exp\{ t\, \sigma^{-1}_{m,n}\, U_{m,n} \}$$

$$E \exp\{ t\, \sigma^{-1}_{m,n}\, \hat{U}_n + t\, \sigma_N^{-1}\, \hat{U}_m \} \exp\{ t\, \sigma^{-1}_{m,n}\, \Delta_{mn} \}$$

$$= E \left(\prod_{j=1}^{N} \exp\{ t\, \sigma^{-1}_{m,n}\, m\, g_2(Y_j) \} \right) \left(\prod_{i=1}^{m} \exp\{ t\, \sigma^{-1}_{m,n}\, n\, g_1(X_i) \} \right)$$

$$\times \left(\exp\{ t\, \sigma^{-1}_{m,n} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi(X_i, Y_j) \} \right)$$

$$= \left(\prod_{j=1}^{n} E \exp \{ t \sigma_{m,n}^{-1} m g_2 (Y_j) \} \right) \left(\prod_{i=1}^{m} E \exp \{ t \sigma_{m,n}^{-1} n g_1 (X_i) \} \right)$$

$$+ E \left(\prod_{j=1}^{n} exp \{ t \sigma_{m,n}^{-1} m g_2 (Y_j) \} \right) \left(\prod_{i=1}^{m} exp \{ t \sigma_{m,n}^{-1} n g_1 (X_i) \} \right)$$

$$\times \left(t \sigma_{m,n}^{-1} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi (X_i, Y_j) \right)$$

$$+ E \left(\prod_{j=1}^{n} \exp \{ t \sigma_{m,n}^{-1} m g_2 (Y_j) \} \right) \left(\prod_{i=1}^{m} \exp \{ t \sigma_{m,n}^{-1} n g_1(X_i) \} \right)$$

$$\times \left(\frac{t^2}{2} \sigma_{m,n}^{-2} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \psi (X_i, Yj) \right)^2 \right)$$

$$+ O (E \mid \sigma_{m,n}^{-1} \Delta_{mn} \mid^{2+\delta})$$

for any $\delta \in (0,1]$.

$$= \gamma_m^n (t) \cdot \gamma_n^m (t)$$

$$+ \gamma_m^{n-1} (t) \cdot \gamma_n^{m-1} (t) \cdot t \sigma_{m,n}^{-1} \binom{n}{1} \binom{m}{1} \times$$

$$E \exp \left\{ t \sigma_{m,n}^{-1} (m g_2 (Y_1) + n g_1 (X_1)) \right\} \psi (X_1, Y_1)$$

(4.2)

$$+ E\left(\prod_{j=1}^{n} \exp\left\{ t\, \sigma_{m,n}^{-1}\, m\, g_2(Y_j) \right\} \right)\left(\prod_{i=1}^{m} \exp\left\{ t\, \sigma_{m,n}^{-1}\, n\, g_1(X_i) \right\} \right)$$

$$\times \left[\frac{t^2}{2}\, \sigma_{m,n}^{-2} \left(\sum_{i,j} \psi^2(X_i, Y_j) + \sum_{i \neq i',\, j} \psi(X_i, Y_j)\, \psi(X_{i'}, Y_j) \right.\right.$$

$$+ \sum_{i,\, j \neq j'} \psi(X_i, Y_j)\, \psi(X_i, Y_{j'})$$

$$\left.\left. + \sum_{i \neq i',\, j \neq j'} \psi(X_i, Y_j)\, \psi(X_{i'}, Y_{j'}) \right) \right]$$

$$+ O\left(E \mid t\, \sigma_{m,n}^{-1}\, \Delta_{mn} \mid^{2+\delta} \right)$$

where $\sigma_{m,n} = O\left(n^{-\frac{3}{2}}\right)$ *when* $m = cn$ *and* $E \mid \Delta_{mn} \mid^{2+\delta} = O(n^2$

Note:

$$E\left[\left(\exp\left(t\, \sigma_{m,n}^{-1}\, m\, g_2(Y_1) \right) - 1 - t\, \sigma_{m,n}^{-1}\, m\, g_2(Y_1) \right) \times \right.$$

$$\left. \left(\exp\left(t\, \sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) - 1 - t\, \sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) \right.$$

$$+ t\, \sigma_{m,n}^{-1}\, m \left(\exp\left(\frac{t\, n\, g_1(X_1)}{\sigma_{m,n}} \right) \right. \tag{4.3}$$

$$- \sum_{\nu=0}^{2} \left(t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) \right)^{\nu} \frac{1}{\nu\,!} \Bigg) g_2 (Y_1)$$

$$+ t\, \sigma_{m,n}^{-1}\, n \left(\exp \left(\frac{t\, m\, g_2 (Y_1)}{\sigma_{m,n}} \right) \right.$$

$$- \sum_{\nu=0}^{2} \left(t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) \right)^{\nu} \frac{1}{\nu\,!} \Bigg) g_1 (X_1)$$

$$+ t^2\, \sigma_{m,n}^{-2}\, n\, m\, g_1 (X_1)\, g_2 (Y_1)$$

$$+ \frac{1}{2}\, t^3\, \sigma_{m,n}^{-3}\, n^2\, m\, g_1^2 (X_1)\, g_2 (Y_1)$$

$$+ \frac{1}{2}\, t^3\, \sigma_{m,n}^{-3}\, n\, m^2\, g_1 (X_1)\, g_2^2 (Y_1) \Bigg] \psi (X_1, Y_1)$$

$$= E \Big[\left(\exp \left(t\, \sigma_{m,n}^{-1} \left(m\, g_2 (Y_1) + n\, g_1 (X_1) \right) \right) - \exp \left(t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) \right) \right)$$

$$- t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) \exp \left(t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) \right) - \exp \left(t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) \right)$$

$$+ 1 + t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) - t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) \exp \left(t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) \right)$$

$$+ t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) + t^2\, \sigma_{m,n}^{-2}\, m\, n\, g_2 (Y_1)\, g_1 (X_1) \Big)$$

$$+ \left(\frac{t\, m\, g_2 (Y_1)}{\sigma_{m,n}} \exp \left(t\, \sigma_{m,n}^{-1}\, n\, g_1 (X_1) \right) - t\, \sigma_{m,n}^{-1}\, m\, g_2 (Y_1) \right.$$

$$- t^2\,\sigma_{m,n}^{-2}\,m\,n\,g_1(X_1)\,g_2(Y_1) - \frac{1}{2}\,t^3\,\sigma_{m,n}^{-3}\,m\,n^2\,g_1^2(X_1)\,g_2(Y_1)\Big)$$

$$+\left(\frac{t\,n\,g_1(X_1)}{\sigma_{m,n}}\exp\left(t\,\sigma_{m,n}^{-1}\,m\,g_2(Y_1)\right) - t\,\sigma_{m,n}^{-1}\,n\,g_1(X_1)\right.$$

$$- t^2\,\sigma_{m,n}^{-2}\,m\,n\,g_1(X_1)\,g_2(Y_1) - \frac{1}{2}\,t^3\,\sigma_{m,n}^{-3}\,m^2\,n\,g_1(X_1)\,g_2^2(Y_1)\Big)$$

$$+ t^2\,\sigma_{m,n}^{-2}\,n\,m\,g_1(X_1)\,g_2(Y_1)$$

$$+ \frac{1}{2}\,t^3\,\sigma_{m,n}^{-3}\,n^2\,m\,g_1^2(X_1)\,g_2(Y_1)$$

$$+ 1/2\,t^3\,\sigma_{m,n}^{-3}\,n\,m^2\,g_1(X_1)\,g_2^2(Y_1)\Big]\,\psi(X_1, Y_1)$$

$$= E\exp\left(t\,\sigma_{m,n}^{-1}\left(m\,g_2(Y_1) + n\,g_1(X_1)\right)\right)\psi(X_1, Y_1)$$

$$- E\exp\left(t\,\sigma_{m,n}^{-1}\,m\,g_2(Y_1)\right)\psi(X_1, Y_1)$$

$$- E\exp\left(t\,\sigma_{m,n}^{-1}\,n\,g_1(X_1)\right)\psi(X_1, Y_1) + E\,\psi(X_1, Y_1)$$

$$= E\exp\left(t\,\sigma_{m,n}^{-1}\left(m\,g_2(Y_1) + n\,g_1(X_1)\right)\right)\psi(X_1, Y_1)$$

which is the last factor in second term in equation (4.2).

Also note that the first term in equation (4.3) is a high order term, i.e.

$$E\left(\exp\left(t\,\sigma_{m,n}^{-1}\,m\,g_2(Y_1)\right) - 1 - \sigma_{m,n}^{-1}\,m\,g_2(Y_1)\right)\times$$

$$\left(\exp\left(t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) - 1 - t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) \cdot \psi(X_1, Y_1)$$

$$= E\Big[\exp\left(t\,\sigma_{m,n}^{-1} \left(m\, g_2(Y_1) + n\, g_1(X_1) \right) \right) - \exp\left(t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1) \right)$$

$$- t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \exp\left(t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1) \right)$$

$$- \exp\left(t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) + 1 + t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1)$$

$$- t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1) \exp\left(t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \right) + t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1)$$

$$+ t^2\,\sigma_{m,n}^{-2}\, m\, n\, g_2(Y_1)\, g_1(X_1) \Big] \times \psi(X_1, Y_1)$$

$$= E\Big[\left(1 + t\,\sigma_{m,n}^{-1} \left(m\, g_2(Y_1) + n\, g_1(X_1) \right) \right.$$

$$+ \frac{t^2\,\sigma_N^{-2}}{2} \left(m\, g_2(Y_1) + n\, g_1(X_1) \right)^2$$

$$\left. + 1/6 \cdot\ t^3\,\sigma_{m,n}^{-3} \left(m\, g_2(Y_1) + n\, g_1(X_1) \right)^3 \right)$$

$$- \left(1 + t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1) + \frac{t^2\,\sigma_{m,n}^{-2}\, m^2\, g_2^2(Y_1)}{2} + \frac{t^3\,\sigma_{m,n}^{-3}\, m^3\, g_2^3(Y_1)}{6} \right)$$

$$- t\,\sigma_{m,n}^{-1}\, n\, g_1(X_1) \left(1 + t\,\sigma_{m,n}^{-1}\, m\, g_2(Y_1) + \frac{t^2\,\sigma_{m,n}^{-2}\, m^2\, g_2^2(Y_1)}{2} \right)$$

$$- \left(1 + t\,\sigma_{m,n}^{-1}\,n\,g_1(X_1) + \frac{t^2\,\sigma_{m,n}^{-2}\,n^2\,g_1^2(X_1)}{2} + \frac{t^3\,\sigma_{m,n}^{-3}\,n^3\,g_1^3(X_1)}{6} \right)$$

$$+ 1 + t\,\sigma_{m,n}^{-1}\,n\,g_1(X_1) + t\,\sigma_{m,n}^{-1}\,m\,g_2(Y_1) - t^2\,\sigma_{m,n}^{-2}\,m\,n\,g_2(Y_1)\,g_1(X_1)$$

$$- t\,\sigma_{m,n}^{-1}\,m\,g_2(Y_1) \left(1 + t\,\sigma_{m,n}^{-1}\,n\,g_1(X_1) + \frac{t^2\,\sigma_{m,n}^{-2}\,n^2\,g_1^2(X_1)}{2} \right)$$

$$+ O\left(n^{-2}\,t^2\right) \Big] \cdot \psi(X_1, Y_1)$$

where we assume that there exists a constant c (greater than zero) such that m = c·n.

$$= E \Bigg[- \frac{2\,t^2\,\sigma_{m,n}^{-2}\,n^2\,c\,g_2(Y_1)\,g_1(X_1)}{2}$$

$$- \frac{3\,t^3\,\sigma_{m,n}^{-3}\,n^2\,c\,g_1(X_1)\,g_2(Y_1)\left(c\,n\,g_2(Y_1) + n\,g_1(X_1)\right)}{6}$$

$$+ t^2\,\sigma_{m,n}^{-2}\,n^2\,c\,g_1(X_1)\,g_2(Y_1) + \frac{t^3\,\sigma_{m,n}^{-3}\,c^2\,n^3\,g_2^2(Y_1)\,g_1(X_1)}{2}$$

$$- t^2\,\sigma_{m,n}^{-2}\,n^2\,c\,g_2(Y_1)\,g_1(X_1)$$

$$+ t^2\,\sigma_{m,n}^{-2}\,n^2\,c\,g_1(X_1)\,g_2(Y_1)$$

$$+ \frac{t^3 \sigma_{m,n}^{-3} n^3 c\, g_1^2(X_1)\, g_2(Y_1)}{2} \Bigg] \cdot \psi(X_1, Y_1) + O(n^{-2} t^4)$$

$$= O(n^{-2} t^4) \tag{4.4}$$

Finally also observe that the second term in equation (4.3) is also negligible. This can be shown as follows.

$$E \Bigg| 2 t \sigma_{m,n}^{-1} m \Bigg[\exp\left(\frac{t n g_1(X_1)}{\sigma_{m,n}} \right) - \sum_{\nu=0}^{2} \frac{\left(t \sigma_{m,n}^{-1} n g_1(X_1) \right)^{\nu}}{\nu !} \Bigg] g_2(Y_1) \cdot \psi(X_1, Y_1) \Bigg|$$

$$\leq 2 \mid t \sigma_{m,n}^{-1} \mid^{3+\theta} n^{2+\theta} m E \mid g_1(X_1) \mid^{2+\theta} \mid \psi(X_1, Y_1)\, g_2(Y_1) \mid$$

$$\forall \quad \theta \in [0,1]$$

If

$$\theta = 6\epsilon = \frac{2\delta}{2+\delta} \text{ and } m = n \cdot c \text{ for some constant c, then we have}$$

$$= 2 \mid t \sigma_{m,n}^{-1} n \mid^{3+6\epsilon} c' E \mid g_1(X_1) \mid^{\frac{4+\delta}{2+\delta}} \mid \psi(X_1, Y_1)\, g_2(Y_1) \mid$$

$$\forall \quad \theta \in [0,1]$$

$$\leq 2 \mid t\, \sigma_{m,n}^{-1}\, n \mid^{3+6\epsilon} c' \left[E\, g_1^4(X_1) \cdot E \mid g_2(Y_1) \mid^{\frac{2+\delta}{1+\delta}} \right]^{\frac{1+\delta}{2+\delta}} \times$$

$$\left[E \mid \psi(X_1, Y_1) \mid^{2+\delta} \right]^{\frac{1}{2+\delta}}$$

By Holder's Inequality.

$$= O\left(\mid n^{-\frac{1}{2}}\, t \mid^{3(1+2\epsilon)} \right) \tag{4.5}$$

$$\left[\because \sigma_N = O\left(n^{-\frac{3}{2}} \right) \text{ when } m = c \cdot n \right]$$

Similarly, we can show that

$$E \left| 2\, t\, \sigma_{m,n}^{-1}\, n \left[\exp\left(\frac{t\, m\, g_2(Y_1)}{\sigma_{m,n}} \right) \right. \right.$$

$$\left. \left. - \sum_{\nu=0}^{2} \frac{\left(t\, \sigma_{m,n}^{-1}\, m\, g_2(Y_1) \right)^{\nu}}{\nu\,!} \right] g_1(X_1) \cdot \psi(X_1, Y_1) \right|$$

$$\leq O\left(\mid n^{-\frac{1}{2}}\, t \mid^{3(1+2\epsilon)} \right) \tag{4.6}$$

Now let us consider the third and final term in equation (4.2).

$$E\left(\prod_{j=1}^{n} \exp\left(t\, \sigma_{m,n}^{-1}\, m\, g_2(Y_j) \right) \right) \cdot \left(\prod_{i=1}^{m} \exp\left(t\, \sigma_{m,n}^{-1}\, n\, g_1(X_i) \right) \right)$$

$$\times \left[\frac{t^2}{2} \sigma_{m,n}^{-2} \left(\sum_{i,j} \psi^2(X_i, Y_j) + \sum_{i \neq i',\, j} \psi(X_i, Y_j)\, \psi(X_{i'}, Y_j) \right.\right.$$

$$+ \sum_{i,\, j \neq j'} \psi(X_i, Y_j)\, \psi(X_i, Y_{j'})$$

$$\left.\left. + \sum_{i \neq i',\, j \neq j'} \psi(X_i, Y_j)\, \psi(X_{i'}, Y_{j'}) \right) \right]$$

$$= \frac{1}{2} \gamma_m^{n-1}(t) \cdot \gamma_n^{m-1}(t) \cdot t^2 \sigma_{m,n}^{-2} m n E \psi^2(X_1, Y_1)$$

$$+ \frac{1}{2} \gamma_m^{n-1}(t) \cdot \gamma_n^{m-2}(t) \cdot (t \sigma_{m,n}^{-1})^4 n^2 2 \binom{m}{2} \binom{n}{1} \times$$

$$E g_1(X_1)\, g_1(X_2)\, \psi(X_1, Y_1)\, \psi(X_2, Y_1)$$

$$+ \frac{1}{2} \gamma_m^{n-2}(t) \cdot \gamma_n^{m-1}(t) \cdot (t \sigma_{m,n}^{-1})^4 m^2 2 \binom{n}{2} \binom{m}{1} \times$$

$$E g_2(Y_1)\, g_2(Y_2)\, \psi(X_1, Y_1)\, \psi(X_1, Y_2)$$

$$+ 4 \cdot \frac{1}{2} \gamma_m^{n-2}(t) \cdot \gamma_n^{m-2}(t) \cdot (t \sigma_{m,n}^{-1})^6 m^2 n^2 \binom{n}{2} \binom{m}{2} \times$$

$$[E g_1(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1)]^2$$

$$+ O(n^{-2} t^4)$$

where m = n·c.

$$= \frac{1}{2}\,\gamma_m^{n-1}(t)\cdot\gamma_n^{m-1}(t)\cdot t^2\,\sigma_{m,n}^{-2}\,m\,n\,E\,\psi^2(X_1, Y_1)$$

$$+\frac{1}{2}\,\gamma_m^{n-1}(t)\,\gamma_n^{m-2}(t)\,t^4\,\sigma_{m,n}^{-4}\,n^2\,m\,(m-1)\,n\,\times$$

$$E\,g_1(X_2)\,g_1(X_1)\,\psi(X_1, Y_1)\,\psi(X_2, Y_1)$$

$$+\frac{1}{2}\,\gamma_m^{n-2}(t)\,\gamma_n^{m-1}(t)\,t^4\,\sigma_{m,n}^{-4}\,m^2\,n\,(n-1)\,m\,\times$$

$$E\,g_2(Y_1)\,g_2(Y_2)\,\psi(X_1, Y_1)\,\psi(X_1, Y_2)$$

$$+\frac{1}{2}\,\gamma_m^{n-2}(t)\,\gamma_n^{m-2}(t)\,t^6\,\sigma_{m,n}^{-6}\,m^2\,n^2\,n\,(n-1)\,m\,(m-1)\,\times$$

$$\left[E\,g_1(X_1)\,g_2(Y_1)\,\psi(X_1, Y_1)\right]^2$$

$$+\ O(n^{-2}\,t^4) \qquad (4.7)$$

where m = n·c.

Thus from equations (4.2) - (4.7), we can say,

$$\phi_N(t) = \gamma_m^n(t)\cdot\gamma_n^m(t)$$

$$+\ \gamma_m^{n-1}(t)\cdot\gamma_n^{m-1}(t)\binom{n}{1}\binom{m}{1}\times$$

$$\left[\frac{t^3\, n\, m}{\sigma_{m,n}^3}\, E\, g_1(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1) \right.$$

$$+ \frac{1}{2}\, t^4\, \sigma_{m,n}^{-4}\, n^2\, m\, E\, g_1^2(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1)$$

$$+ \frac{1}{2}\, t^4\, \sigma_{m,n}^{-4}\, n\, m^2\, E\, g_1(X_1)\, g_2^2(Y_1)\, \psi(X_1, Y_1)$$

$$\left. + \frac{1}{2}\, \frac{t^2}{\sigma_{m,n}^2}\, E\, \psi^2(X_1, Y_1) \right]$$

$$+ \frac{1}{2}\, \gamma_m^{n-1}(t) \cdot \gamma_n^{m-2}(t)\, t^4\, \sigma_{m,n}^{-4}\, n^3\, m\, (m-1) \times$$

$$E\, g_1(X_1)\, g_1(X_2)\, \psi(X_1, Y_1)\, \psi(X_2, Y_1)$$

$$+ \frac{1}{2}\, \gamma_m^{n-2}(t) \cdot \gamma_n^{m-1}(t)\, t^4\, \sigma_{m,n}^{-4}\, m^3\, n\, (n-1) \times$$

$$E\, g_2(Y_1)\, g_2(Y_2)\, \psi(X_1, Y_1)\, \psi(X_1, Y_2)$$

$$+ \frac{1}{2}\, \gamma_m^{n-2}(t) \cdot \gamma_n^{m-2}(t) \cdot \frac{t^6\, m^3\, n^3}{\sigma_{m,n}^6}\, (n-1)\, (m-1) \times$$

$$\left[E\, g_1(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1) \right]^2 \qquad (4.8)$$

$$+ \; O\left(\mid \gamma_m^{n-2}(t) \cdot \gamma_n^{m-2}(t) \mid \cdot P(\mid t \mid) n^{-1-3\epsilon} + \mid n^{-\frac{1}{2}} t \mid^{2+\delta} \right)$$

where P is a polynomial.

4.2 Two Sample: Degree (K_1, K_2)

Let $X_1, X_2, \ldots, X_m$ be independently and identically distributed (i.i.d.) random variables with distribution function F(X) and let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with distribution function G(Y). If h is a function of $K_1 + K_2$ arguments,

$$h(X_{i_1}, \ldots, X_{i_{k_1}}, Y_{j_1}, \ldots, Y_{j_{k_2}})$$

which is symmetric in its K_1 X's and K_2 Y's, then a generalized two sample U-statistic $(U_{m,n})$ based on a symmetric kernel h of degree K_1 and K_2 is defined as follows:

If $\underset{\sim}{X}_i = (X_{i_1}, \ldots, X_{i_{k_1}})$ and $\underset{\sim}{Y}_j = (Y_{j_1}, \ldots, Y_{j_{k_2}})$ then

$$U_{m,n} = \sum_{1 \le i_1 < \ldots < i_{k_1} \le m \,;\, 1 \le j_1 < \ldots < j_{k_2} \le n} h(\underset{\sim}{X}_i, \underset{\sim}{Y}_j) \,. \tag{4.9}$$

Note that again we consider a standardized version of two sample U-statistics of degree K_1 and K_2. We can assume without loss of generality that

$$E\, h(\underset{\sim}{X}, \underset{\sim}{Y}) = 0 \quad \text{and} \quad E\, h^2(\underset{\sim}{X}, \underset{\sim}{Y}) < \infty \,.$$

where $\underset{\sim}{X} = (X_1, ..., X_{K_1})$ and $\underset{\sim}{Y} = (Y_1, ..., Y_{K_2})$.

Let us define

$$g_{i,j}(x_1, ..., x_i, y_1, ..., y_j)$$

$$= E(h(\underset{\sim}{X}, \underset{\sim}{Y}) \mid X_1 = x_1, ..., X_i = x_i, Y_1 = y_1, ..., Y_j = y_j)$$

where $i = 0,1, ..., K_1$ & $j = 0,1, ..., K_2$.

Now define,

$$\psi_{0,1}(Y_1) = E(h(\underset{\sim}{X}, \underset{\sim}{Y}) \mid Y_1) = g_{0,1}(Y_1)$$

$$\psi_{1,0}(X_1) = E(h(\underset{\sim}{X}, \underset{\sim}{Y}) \mid X_1) = g_{1,0}(X_1)$$

$$\psi_{1,1}(X_1, Y_1) = g_{1,1}(X_1, Y_1) - \psi_{0,1}(Y_1) - \psi_{1,0}(X_1)$$

$$= g_{1,1}(X_1, Y_1)$$

$$- \sum_{j+j'=1} \sum_{1,j:1,j'} \psi_{j,j'}(X_i, ..., X_{i_j}, Y_{i'_1}, ..., Y_{i'_{j'}})$$

where i is used as an index for X, i´ is used as an index for Y and the double sum is over all the number of distinct combinations of size j(=0 or 1) out of one X and also over all the number of distinct combinations of size j´(=0 or 1), out of one Y.

Similarly,

$$\psi_{1,2}(X_1, Y_1, Y_2) = g_{1,2}(X_1, Y_1, Y_2) - \sum_{j+j'=1}^{2} \sum_{1,j:2,j'} \psi_{j,j'}(X_i, ..., X_{i_j}, Y_{i'_1}, ..., Y_{i'_{j'}})$$

$$= g_{1,2}(X_1, Y_1, Y_2) - [\psi_{1,1}(X_1, Y_1) + \psi_{1,1}(X_1, Y_2) + \psi_{0,2}(Y_1, Y_2)] - [\psi_{0,1}(Y_1) + \psi_{1,0}(X_1) + \psi_{0,1}(Y_2)]$$

$$\psi_{2,1}(X_1, X_2, Y_1) = g_{2,1}(X_1, X_2, Y_1) - \sum_{j+j'=1}^{2} \sum_{2,j:1,j'} \psi_{j,j'}(X_i, ..., X_{i_j} Y_{i'_1}, ..., Y_{i'_{j'}})$$

$$= g_{2,1}(X_1, X_2, Y_1) - [\psi_{1,1}(X_1, Y_1) + \psi_{1,1}(X_2, Y_1) + \psi_{2,0}(X_1, X_2)] - [\psi_{1,0}(X_1) + \psi_{1,0}(X_2) + \psi_{0,1}(Y_2)]$$

. .

. .

. .

$$\psi_{r,s}(X_1, .., X_r, Y_1, .., Y_s) = g_{r,s}(X_1, .. X_r, Y_1, .., Y_s)$$

$$- \sum_{j+j'=1}^{r+s-1} \sum_{r,j:s,j'} \psi_{j,j'} (X_{i_1}, ..., X_{i_j} Y_{i'_1}, ..., Y_{i'_{j'}})$$

where i is used as an index for X, i′ is used as an index for Y and the double sum is over all the number of distinct combinations of size j(=0,1, ..., r), out of a total of r = 0,1, ..., K_1, X's and also over all the number of distinct combinations of size j′(=0,1, ..., s) out of a total of s = 0,1, ..., K_2, Y's.

. .

. .

. .

$$\psi_{k_1,k_2}(X_1, .., X_{k_1}, Y_1, .., Y_{k_2}) = g_{k_1,k_2}(X_1, .., X_{k_1}, Y_1, .., Y_{k_2})$$

$$- \sum_{j+j'=1}^{k_1+k_2-1} \sum_{k_1,j:k_2,j'} \psi_{j,j'} (X_i, ..., X_{ij}, Y_{i'_1}, ..., Y_{i'_{j'}})$$

Also as we defined in Chapter 3, let

$$\zeta_{n-j,k-j} = \frac{(n-j)!}{(n-k)!} = (n-j)(n-2)\ldots(n-k+1) = O(n^{k-j})$$

then we further define,

$$U_{1,0} = \zeta_{m-1,k_1-1} \zeta_{n,k_2} \sum_{i=1}^{m} \psi_{1,0} (X_i)$$

$$(m-1)...(m-k_1+1) \cdot n...(n-k_2+1) \sum_{i=1}^{m} \psi_{1,0}(X_i)$$

$$U_{0,1} = \zeta_{m,k_1} \zeta_{n-1,k_2-1} \sum_{j=1}^{n} \psi_{0,1}(Y_j)$$

$$= m...(m-k_1+1) \cdot (n-1)...(n-k_2+1) \sum_{j=1}^{n} \psi_{0,1}(Y_j)$$

$$U_{1,1} = \zeta_{m-1,k_1-1} \zeta_{n-1,k_2-1} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi_{1,1}(X_i, Y_j)$$

$$= (m-1)...(m-k_1+1) \cdot (n-1)...(n-k_2+1) \sum_{i=1}^{m} \sum_{j=1}^{n} \psi_{1,1}(X_i, Y_j)$$

. .

. .

. .

$$U_{r,s} = \zeta_{m-r,k_1-r} \zeta_{n-s,k_2-s} \times$$

$$\sum_{\substack{1 \le i_1 < \ldots < i_r \le m \\ 1 \le j_1 < \ldots < j_s \le n}} \cdots \sum \psi_{r,s}(X_{i_1}, \ldots, X_{i_r}, Y_{j_1}, \ldots, Y_{j_s})$$

. .

. .

$$U_{k_1,\,k_2} = \zeta_{m,0}\,\zeta_{n,0} \sum_{\substack{1 \le i_1 < \ldots < i_{k_1} \le m \\ 1 \le j_1 < \ldots < j_{k_2} \le n}} \cdots \sum \psi_{k_1,\,k_2}\,(X_{i_1}, \ldots X_{i_{k_1}}, Y_{j_1}, \ldots, Y_{j_{k_2}})$$

where $\zeta_{m,0} = \zeta_{n,0} = 1$.

Then, note that

$$U_{m,n} = \sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \frac{U_{r,s}}{(k_1-r)!\,(k_2-s)!} \tag{4.10}$$

Where $U_{0,0} = 0$.

Proof of the above result, Hoeffding's decomposition in the two sample case is similar to the one sample proof given on page 19 of this manuscript. A. J. Lee (1990), in his book "U-Statistics", Theorem 3, page 40, also states the same result in a slightly different set up.

By using the independence of X_i's and Y_j's, together with the result of Lemma 3.2. in Chapter 3, it implies that

$\psi_{r,s}\,(X_{i_1}, \ldots, X_{i_r}, Y_{j_1}, \ldots Y_{j_s})$ and

$$\psi_{r',s'}\,(X_{i'_1}, \ldots, X_{i'_{r'}}, Y_{j'_1}, \ldots Y_{j'_{s'}}) \tag{4.11}$$

for $[(i_1, i_2, \ldots, i_r), (j_1, j_2, \ldots, j_s)] \neq [(i'_1, i'_2, \ldots, i'_{r'}), (j'_1, j'_2 \ldots, j'_{s'})]$, ie,

at least one of two groups of subscripts are different, are uncorrelated.

Thus,

$$\sigma^2_{m,n} = Var\ (\ U_{m,n}\ ) = \sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \frac{Var\ (\ U_{r,s}\ )}{[\ (\ k_1-r)\ !\cdot (\ k_2-s)\ !\ ]^2}$$

$$= \sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \left[\frac{m!}{(m-k_1)!} \frac{(m-r)!}{m!} \frac{1}{(k_1-r)!} \frac{n!}{(n-k_2)!} \frac{(n-s)!}{n!} \times \right.$$

$$\left. \frac{1}{(k_2-s)\ !} \right]^2 \binom{m}{r} \binom{n}{s} E\ \psi^2_{r,s}\ (X_1, \ldots, X_r, Y_1, \ldots, Y_s\ )$$

$$= \sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \left[\frac{(m-r)!}{(m-k_1)!\ (k_1-r)!} \cdot \frac{(n-s)!}{(n-k_2)!\ (k_2-s)!} \right]^2 \binom{m}{r} \binom{n}{s} \times$$

$$E\ \psi^2_{r,s}(\ X_1, \ldots\ X_r, Y_1, \ldots\ Y_s\ )$$

$$= \sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \left[\binom{(m-r)}{(k_1-r)} \binom{(n-s)}{(k_2-s)} \right]^2 \cdot \binom{m}{r} \binom{n}{s} \times$$

$$E\ \psi^2_{r,s}\ (\ X_1, \ldots, X_r, Y_1, \ldots, Y_s\ ) \qquad (4.12)$$

$\because$ X_i's are all i.i.d., Y_j's are all i.i.d. and furthermore, there is independence between X's and Y's.

By choosing m = c·n (for some c > 0), it is quite obvious that $\sigma^2\ (\ U_{1,0}\ )$ and $\sigma^2\ (\ U_{0,1}\ )$ dominate the right-hand side and also note,

$$\sigma^2_{m,n} = O\left(n^{2(k_1+k_2)-1} \right) \tag{4.13}$$

$$\Leftrightarrow \sigma_{m,n} = O\left(n^{(k_1+k_2)-1/2} \right).$$

Now based on $\lambda_{N-j,k-j}$ (See page 26, Chapter 3) if we define

$$\delta_{m-r,n-s} = \lambda_{m-r,k_1-r} \cdot \lambda_{n-s,k_2-s} \tag{4.14}$$

and then consider,

$$\phi_{m,n}(t) = E \exp\left[t\, \sigma^{-1}_{m,n} \cdot U_{m,n} \right]$$

$$= E \exp\left[t\, \sigma^{-1}_{m,n} \left(\sum_{r=0}^{k_1} \sum_{s=0}^{k_2} \frac{U_{r,s}}{(k_1-r)!\,(k_2-s)!} \right) \right]$$

$$= E\left[\exp\left(t\, \sigma^{-1}_{m,n} \frac{U_{1,0}}{(k_1-1)!\,k_2!} \right) \exp\left(t\, \sigma^{-1}_{m,n} \frac{U_{0,1}}{k_1!\,(k_2-1)!} \right) \times \right.$$

$$\left. \exp\left(\frac{t\, U_{-1}}{\sigma_{m,n}} \right) \right]$$

$$\text{where} \quad U_{-i} = \sum_{r+s>i} \sum \frac{U_{r,s}}{(k_1-r)!\,(k_2-s)!}$$

$$= E\left[\left(\prod_{i=1}^{m} \exp\left[t\, \sigma^{-1}_{m,n}\, \delta_{m-1,n}\, \psi_{1,0}(X_i) \right] \right) \times \right.$$

$$\left(\prod_{j=1}^{n} \exp \left[t\, \sigma_{m,n}^{-1}\, \delta_{m,n-1}\, \psi_{0,1}\, (Y_j) \right] \right) \times$$

$$\left\{ 1 + t\, \sigma_{m,n}^{-1} \left[\left(\frac{U_{1,1}}{(k_1 - 1)!\, (k_2 - 1)!} + \frac{U_{2,0}}{(k_1 - 2)!\, k_2!} + \frac{U_{0,2}}{k_1!\, (k_2 - 2)!} \right) + \right. \right.$$

$$\left(\frac{U_{2,1}}{(k_1 - 2)!\, (k_2 - 1)!} + \frac{U_{1,2}}{(k_1 - 1)!\, (k_2 - 2)!} + \frac{U_{3,0}}{(k_1 - 3)!\, k_2!} + \right.$$

$$\left. + \frac{U_{0,3}}{k_1!\, (k_2 - 3)!} \right) + U_{-3} \Bigg] +$$

$$1/2 \left[t \sigma_{m,n}^{-1} \left(\frac{U_{1,1}}{(k_1 - 1)!\, (k_2 - 1)!} + \frac{U_{2,0}}{(k_1 - 2)!\, k_2!} + \frac{U_{0,2}}{k_1!\, (k_2 - 2)!} \right) + U_{-2} \right]^2 \Bigg\}$$

$$+ O \left(E \mid t\, \sigma_{m,n}^{-1}\, U_{-1} \mid^{2+\delta} \right) \quad \textit{for some} \quad \delta > 0 \tag{4.15}$$

Choosing m = c·n (for some c > 0) and by using lemma 3.4. in Chapter 3 we know that

$$E \mid U_{-j} \mid^{2+\delta} = O \left(n^{(k_1+k_2-(j+1))/2)\,(2+\delta)} \right) \tag{4.16}$$

$$\textit{where} \quad \delta \geq 0.$$

and also,

$$E\left[\left(\prod_{i=1}^{m} \exp\left[t\, \sigma_{m,n}^{-1}\, \delta_{m-1,n}\, \psi_{1,0}\left(X_i \right) \right] \right) \times \right. \tag{4.17}$$

$$\left. \left(\prod_{j=1}^{n} \exp\left[t\, \sigma_{m,n}^{-1}\, \delta_{m,n-1}\, \psi_{0,1}\left(Y_j \right) \right] \right) t\sigma_{m,n}^{-1}\, U_{-3} \right] \;=\; O\left(n^{-\frac{3}{2}} \right).$$

Noting that interaction terms within the square term in the equation (4.15) are also of order $O(n^{-3/2})$, thus we have

$$\phi_N (t) = \gamma_n^m(t) \cdot \gamma_m^n(t)$$

$$+ \gamma_n^{m-1}(t)\, \gamma_m^{n-1}(t)\, (t\, \sigma_{m,n}^{-1})^3 \binom{m}{1} \binom{n}{1} \delta_{m-1,n}\, \delta_{m,n-1}\, \delta_{m-1,n-1} \times$$

$$E\, \psi_{1,0}(X_1)\, \psi_{0,1}(Y_1)\, \psi_{1,1}(X_1, Y_1)$$

$$+ \frac{1}{2}\, \gamma_n^{m-1}(t)\, \gamma_m^{n-1}(t)\, (t\, \sigma_{m,n}^{-1})^4 \binom{m}{1} \binom{n}{1} \delta_{m-1,n}\, \delta_{m,n-1}\, \delta_{m-1,n-1} \times$$

$$\left[\delta_{m-1,n}\, E\, \psi_{1,0}^2(X_1)\, \psi_{0,1}(Y_1)\, \psi_{1,1}(X_1\, Y_1) + \right.$$

$$\left. \delta_{m,n-1}\, E\, \psi_{1,0}(X_1)\, \psi_{0,1}^2(Y_1)\, \psi_{1,1}(X_1, Y_1) \right]$$

$$+ \frac{1}{2}\, \gamma_n^{m-1}(t)\, \gamma_m^{n-1}(t)\, (t\, \sigma_{m,n}^{-1})^2\, m\, n\, \delta_{m-1,n-1}^2 \cdot E\, \psi_{1,1}^2(X_1, Y_1)$$

$$+ \frac{1}{2} \gamma_n^{m-2}(t)\,\gamma_m^{n-1}(t)\,(t\,\sigma_{m,n}^{-1})^4 \cdot \binom{m}{2}\binom{n}{1} \cdot 2\,\delta_{m-1,n}^2\,\delta_{m-1,n-1}^2 \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{1,1}(X_1,Y_1)\,\psi_{1,1}(X_2,Y_1)$$

$$+ \frac{1}{2} \gamma_n^{m-1}(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{1}\binom{n}{2} 2\,\delta_{m,n-1}^2\,\delta_{m-1,n-1}^2 \times$$

$$E\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{1,1}(X_1,Y_1)\,\psi_{1,1}(X_1,Y_2)$$

$$+ \frac{1}{2} \gamma_n^{m-2}(t)\,\gamma_m^{n-2}(t)\,4\,(t\,\sigma_{m,n}^{-1})^6 \binom{m}{2}\binom{n}{2} (\delta_{m-1,n-1}\,\delta_{m-1,n}\,\delta_{m,n-1})^2 \times$$

$$\left[\,E\,\psi_{1,0}(X_1)\,\psi_{0,1}(Y_1)\,\psi_{1,1}(X_1\,Y_1)\,\right]^2$$

$$+ \gamma_n^{m-2}(t)\,\gamma_m^{n-1}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{2}\binom{n}{1} \delta_{m-1,n}^2\,\delta_{m,n-1}\,\delta_{m-2,n-1} \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{0,1}(Y_1)\,\psi_{2,1}(X_1,X_2,Y_1)$$

$$+ \gamma_n^{m-1}(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{1}\binom{n}{2} \delta_{m-1,n}\,\delta_{m,n-1}^2\,\delta_{m-1,n-2} \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{1,2}(X_1,Y_1,Y_2)$$

$$+ \gamma_n^{m-2}(t)\,\gamma_m^{n}(t)\,(t\,\sigma_{m,n}^{-1})^3 \binom{m}{2} \delta_{m-1,n}^2\,\delta_{m-2,n} \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{2,0}(X_1,X_2)$$

$$+ \gamma_n^m(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^3 \binom{n}{2} \delta_{m,n-1}^2\,\delta_{m,n-2} \times$$

$$E\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{0,2}(Y_1, Y_2)$$

$$+ \gamma_n^{m-2}(t)\,\gamma_m^n(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{2} \delta_{m-1,n}^3\,\delta_{m-2,n} \times$$

$$E\,\psi_{1,0}^2(X_1)\,\psi_{1,0}(X_2)\,\psi_{2,0}(X_1, X_2)$$

$$+ \gamma_n^m(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{n}{2} \delta_{m,n-1}^3\,\delta_{m,n-2} \times$$

$$E\,\psi_{0,1}^2(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{0,2}(Y_1, Y_2)$$

$$+ \gamma_n^{m-3}(t)\,\gamma_m^n(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{3} \delta_{m-1,n}^3\,\delta_{m-3,n} \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{1,0}(X_3)\,\psi_{3,0}(X_1, X_2, X_3)$$

$$+ \gamma_n^m(t)\,\gamma_m^{n-3}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{n}{3} \delta_{m,n-1}^3\,\delta_{m,n-3} \times$$

$$E\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{0,1}(Y_3)\,\psi_{0,3}(Y_1, Y_2, Y_3)$$

$$+ \frac{1}{2}\,\gamma_n^{m-2}(t)\,\gamma_m^n(t)\,(t\,\sigma_{m,n}^{-1})^2 \binom{m}{2} \delta_{m-2,n}^2 \cdot E\,\psi_{2,0}^2(X_1, X_2)$$

$$+ \frac{1}{2}\,\gamma_n^m(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^2 \binom{n}{2} \delta_{m,n-2}^2 \cdot E\,\psi_{0,2}^2(Y_1, Y_2)$$

$$+ 3\ \gamma_n^{m-3}(t)\,\gamma_m^{n}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{3}\,\delta_{m-1,n}^2\,\delta_{m-2,n}^2 \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{2,0}(X_1,X_3)\,\psi_{2,0}(X_2,X_3)$$

$$+ 3\ \gamma_n^{m}(t)\,\gamma_m^{n-3}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{n}{3}\,\delta_{m,n-1}^2\,\delta_{m,n-2}^2 \times$$

$$E\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{0,2}(Y_1,Y_3)\,\psi_{0,2}(Y_2,Y_3)$$

$$+ 3\ \gamma_n^{m-4}(t)\,\gamma_m^{n}(t)\,(t\,\sigma_{m,n}^{-1})^6 \binom{m}{4}\left(\delta_{m-1,n}^2\,\delta_{m-2,n}\right)^2 \times$$

$$\left[E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{2,0}(X_1,X_2)\right]^2$$

$$+ 3\ \gamma_n^{m}(t)\,\gamma_m^{n-4}(t)\,(t\,\sigma_{m,n}^{-1})^6 \binom{n}{4}\left(\delta_{m,n-1}^2\,\delta_{m,n-2}\right)^2 \times$$

$$\left[E\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\,\psi_{0,2}(Y_1,Y_2)\right]^2$$

$$+ 2\ \gamma_n^{m-2}(t)\,\gamma_m^{n-1}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{m}{2}\,n\,\delta_{m-1,n}\,\delta_{m,n-1}\,\delta_{m-1,n-1}\,\delta_{m-2,n} \times$$

$$E\,\psi_{1,0}(X_2)\,\psi_{0,1}(Y_1)\,\psi_{1,1}(X_1,Y_1)\,\psi_{2,0}(X_1,X_2)$$

$$+ 2\ \gamma_n^{m-1}(t)\,\gamma_m^{n-2}(t)\,(t\,\sigma_{m,n}^{-1})^4 \binom{n}{2}\,m\,\delta_{m-1,n}\,\delta_{m,n-1}\,\delta_{m-1,n-1}\,\delta_{m,n-2} \times$$

$$E\,\psi_{0,1}(Y_2)\,\psi_{1,0}(X_1)\,\psi_{1,1}(X_1,Y_1)\,\psi_{0,2}(Y_1,Y_2)$$

$$+ \gamma_n^{m-2}(t)\,\gamma_m^{n-2}(t)(t\,\sigma_{m,n}^{-1})^6 \binom{n}{2}\binom{m}{2} \delta_{m-1,n}^2\, \delta_{m,n-1}^2\, \delta_{m-2,n}\, \delta_{m,n-2} \times$$

$$E\,\psi_{1,0}(X_1)\,\psi_{1,0}(X_2)\,\psi_{0,1}(Y_1)\,\psi_{0,1}(Y_2)\cdot$$

$$\psi_{2,0}(X_1,X_2)\,\psi_{0,2}(Y_1,Y_2)$$

$$+ O\left(\left|n^{-\frac{1}{2}}t\right|^{2+\delta}\right) \qquad (4.18)$$

The preceeding approximation (4.18) of cumulant generating function for the general two sample U-statistic is based on a kernel of degree K_1 and K_2. Although this approximation involves a total of 21 expectations, some of these are the same because of their symmetric nature. It should be noted that the approximation (4.8) for cumulant generating function of a two sample U-statistic based on a kernel of degree (1,1) is a special case of equation 4.18 and it involves only six expectations, some of which may also be similar.

4.3 Example

For a two sample U-statistics, we used Mann-Whitney-Wilcoxon Rank Sum test statistic as an example to test the accuracy of the above approximation. Due to the non-existence of a closed-form expression for the cumulant generating function of the Mann-Whitney-Wilcoxon test statistic, we had to approximate the cumulant generating function. We used (after duplication) the null distribution of Mann-Whitney-Wilcoxon statistic (for m=10 and n=5) as published by Buckle, Kraft and

van Eeden: "Tables Prolongeese de la Distribution de Mann-Whitney-Wilcoxon" (1969b) and also adapted in Lehman's (1975) book to approximate the moment generating function and hence the cumulant generating function. This was used to compare not only with our approximation of the cumulant generating function but also the approximation of the cumulant generating function which was calculated by the first four exact cumulants of the Mann-Whitney-Wilcoxon test statistics as provided by Sundrum (1954), see Figure 4.1.

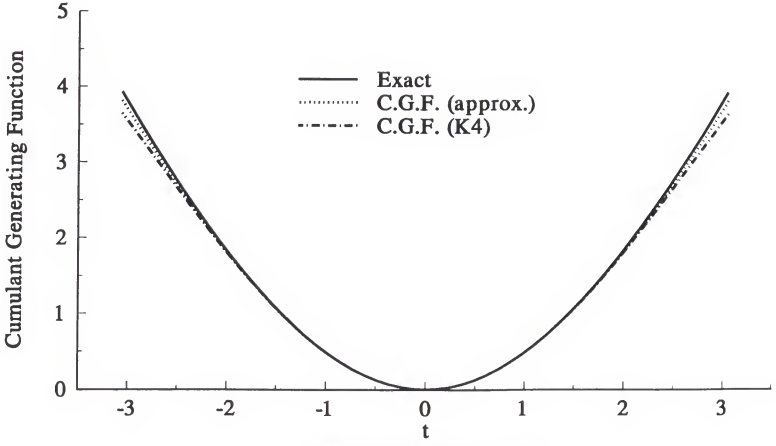
Let $X_1, X_2, ..., X_m$ be independently and identically distributed (i.i.d.) random variables with distribution function F(X) and let $Y_1, Y_2, ..., Y_n$ be i.i.d. random variables with distribution function $G(Y) = F(Y-\Delta)$, where $-\infty < \Delta < \infty$ is an unknown shift parameter. The null hypothesis of interest for this setting is $H_o: \Delta = 0$, and the alternative hypothesis can be any one of the three alternatives, $H_1: \Delta >, <,$ or $\neq 0$. To test the above hypothesis the Mann-Whitney-Wilcoxon test statistic as proposed by Mann and Whitney (1947) is

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(Y_j - X_i) ,$$

where

$$\phi(X) = \begin{cases} 1 & if \ X > 0 \\ 0 & otherwise. \end{cases}$$

Let $E = [\phi(Y_j - X_i)] = p$, then for deriving our approximation for cumulant generating function of Mann-Whitney-Wilcoxon test statistic the corresponding



Mann-Whitney-Wilcoxon Test
Comparison of Cumulant Generating Function
Using Exact, C.G.F. (approx.) and C.G.F. (K4)
Figure 4.1

U-statistics which is centered about zero can be defined by

$$U_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} h(X_i, Y_j)$$

where $h(X_i, Y_j) = \phi(Y_j - X_i) - p$. Note that again we do not follow the usual convention of dividing the sum in the above formulation by the number m·n of its terms. Since our results concern the standardized version of $U_{m,n}$, this does not make any difference. To apply the Hoeffding's decomposition we defined

$$g_1(x) = E\, h((X_i, Y_j) \mid X_i = x)$$

$$= 1 - G(x) - p ,$$

$$g_2(y) = E\, h((X_i, Y_j) \mid Y_j = y)$$

$$= F(y) - p \quad \text{and}$$

$$\psi(x, y) = h(x, y) - g_1(x) - g_2(y)$$

$$= F(y) - p .$$

Under this setup we used the result from the approximation (for moment generating function) as given in equation 4.8 to aproximate the tail probability by the saddlepoint approximation as given in Chapter 2, equation (2.8). To do this we first had to approximate six expectations (of which some were same) under the assumption that the two underlying population distributions are identical, i.e., null hypothesis. Again since U-statistics have a distribution free property, these expectations are independent of the nature of the underlying distribution. These

expectations were found to be as follows:

(i) $E_1 = E\, g_1(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1) = 0$,

(ii) $E_2 = E\, g_1^2(X_1)\, g_2(Y_1)\, \psi(X_1, Y_1) = -0{\cdot}027778$,

(iii) $E_3 = E\, g_1(X_1)\, g_2^2(Y_1)\, \psi(X_1, Y_1) = -0{\cdot}027778$,

(iv) $E_4 = E\, \psi^2(X_1, Y_1) = 0{\cdot}083333$,

(v) $E_5 = E\, g_1(X_1)\, g_1(X_2)\, \psi(X_1, Y_1)\, \psi(X_2, Y_1) = 0{\cdot}013889$,

(vi) $E_6 = E\, g_2(Y_1)\, g_2(Y_2)\, \psi(X_1, Y_1)\, \psi(X_1, Y_2) = 0{\cdot}013889$.

Note that in equation 4.8, except for these expectations, everything else is either a constant or a function of t. Thus when we use eq. 4.8 to approximate the cumulant generating function (of Mann-Whitney test statistics) and henceforth use equation 2.2 in Chapter 2, we come up with an equation in t which we would like to minimize with respect to t. The value of t for which this equation is minimized is the saddlepoint, needed for the saddlepoint approximation.

Besides the normal approximation, we again computed two other saddlepoint approximations.

The first was based on the cumulant generating function as evaluated by the null distribution of the Mann-Whitney-Wilcoxon statistic and, since the tail probability based on the saddlepoint approximation (eq. 2.8) was very close to that of the null distribution of Mann-Whitney-Wilcoxon test statistic (see Table 4.1), this tail probability approximation was again excluded from Figure 4.2 to keep the

Table 4.1. Comparison of Tail probability approximation of Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) between Exact(Null) distribution, Normal approximation and Saddlepoint approximations.

W	W_{std}	Sp_app	Exact	Sp_ex	Sp_K_4	Normal
1	-2.939390		.0007	0.000716267		0.00164431
2	-2.816910		.0013	0.001357470		0.00242438
3	-2.694440		.0023	0.002408450		0.00352536
4	-2.571960		.0040	0.004014450		0.00505617
5	-2.449490		.0063	0.006359880		0.00715294
6	-2.327020		.0097	0.009662700		0.00998223
7	-2.204540	0.0159425	.0140	0.014171300		0.01374320
8	-2.082070	0.0209580	.0200	0.201600000		0.01866820
9	-1.959590	0.0283532	.0276	0.027921300		0.02502180
10	-1.837120	0.0379989	.0376	0.037757500		0.03309630
11	-1.714640	0.0501082	.0496	0.049968600		0.04320540
12	-1.592170	0.0649349	.0646	0.064839400	0.0735473	0.05567734
13	-1.469690	0.0827180	.0823	0.082625000	0.0884643	0.07082230
14	-1.347220	0.1036560	.1032	0.103536000	0.1078120	0.08895480
15	-1.224740	0.1278910	.1272	0.127725000	0.1309280	0.11033600
16	-1.102270	0.1554920	.1548	0.155273000	0.1576530	0.13517200
17	-0.979796	0.1864470	.1855	0.186178000	0.1878990	0.16359300
18	-0.857321	0.2206560	.2198	0.220347000	0.2215380	0.19563400
19	-0.734847	0.2579250	.2567	0.257596000	0.2583700	0.23121600
20	-0.612372	0.2979740	.2970	0.297649000	0.2981080	0.27014600
21	-0.489898	0.3404200	.3393	0.340122000	0.3403610	0.31210300
22	-0.367423	0.3848270	.3839	0.384580000	0.3846810	0.35665200
23	-0.244949	0.4306760	.4296	0.430501000	0.4305300	0.40324800
24	-0.122474		.4765	0.477318000		0.45126200

Where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test-Statistic

Sp_app = Saddlepoint Approximation Based on Our Approximation

Exact = Null Distribution of Test-Statistic

Sp_ex = Saddlepoint Approximation Based on Exact C.G.F.

Sp_K_4 = Saddlepoint Approximation Based on Exact Cumulants (K_1-K_4)

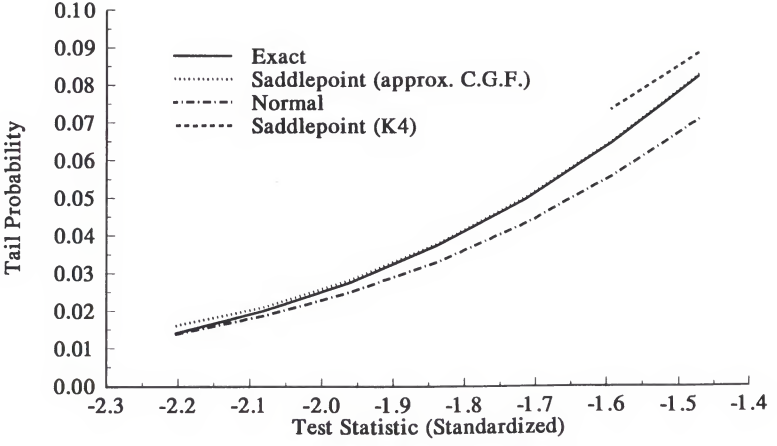
Normal = Normal Approximation

number of comparisons optimal.

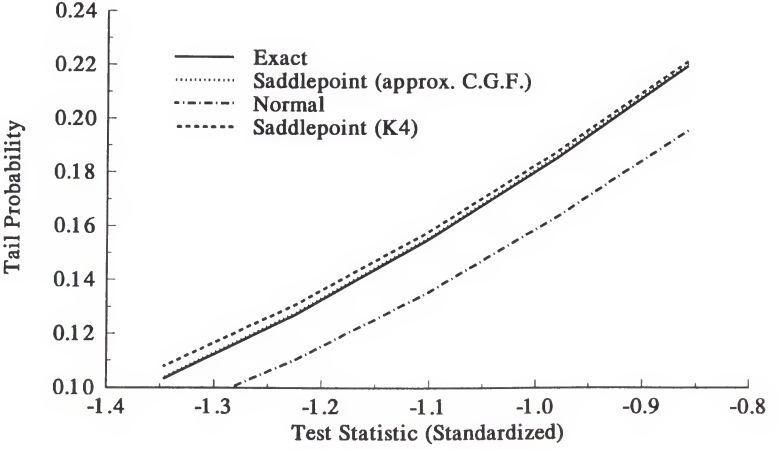
For the second Saddlepoint approximation, we used an approach of G. S. Easton and E. Ronchetti (1986) where they proposed to approximate the first four cumulants of the statistic and in turn used these to further approximate the cumulant generating function used in the saddlepoint approximation. Here, mainly to optimize their approximation, we used the exact (first four) cumulants of the Mann-Whitney-Wilcoxon statistic as given by Sundaram (1954) to approximate the cumulant generating function used in the saddlepoint approximation.

Comparisons of these approximations with the exact have been tabulated in Table 4.1 and plotted in Figures 4.2(A-C). From Table 4.1, it is clear that the Saddlepoint Approximation works quite well to approximate the tail probability under the null distribution for the Mann-Whitney-Wilcoxon test statistic. In fact it is quite close (up to the fourth decimal place) to the exact if cumulant generating function calculated by null distribution itself is used, see Table 4.1. While saddlepoint approximation based on our approximation of the cumulant generating function also works quite well. It not only beats the normal (large sample) approximation over almost the whole range, but it also beats the saddlepoint approximation as proposed by Easton and Ronchetti (1986) in JASA where the cumulant generating function has been approximated by using the first four cumulants (exact).

For an extreme test statistic value, when the tail probability becomes very small, i.e., the standardized test statistic becomes quite large, the saddlepoint also



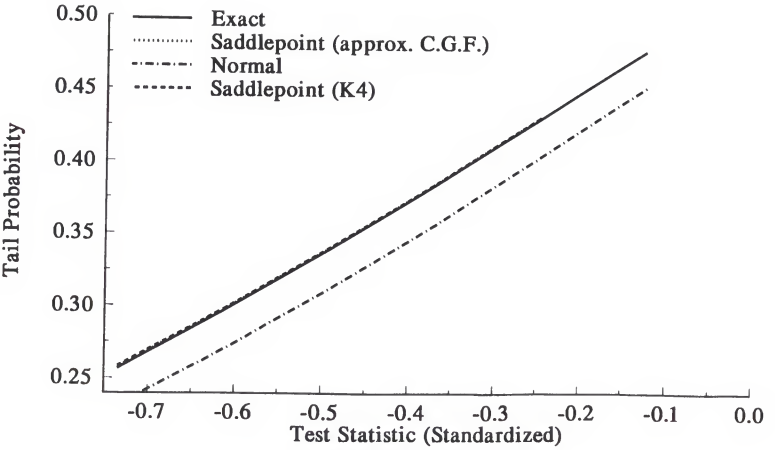
Mann-Whitney-Wilcoxon Test
Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution
Figure 4.2 (A)



Mann-Whitney-Wilcoxon Test

Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution

Figure 4.2 (B)



Mann-Whitney-Wilcoxon Test
Comparison of Tail Probability Approximation
Using Exact, Saddle-Point and Normal Distribution
Figure 4.2 (C)

becomes relatively larger and then the saddlepoint approximation (based on the cumulant generating function based on Hoeffding's decomposition) tends to lose itsaccuracy. The reason for this is that for large values of t, cumulant generating function approximated by the decomposition theorem may not be too accurate and hence the saddlepoint itself may be less accurate. If this becomes a problem, we suggest that some of the higher order terms may again need to be considered. It may be worthwhile to note that this is not a concern for saddlepoint approximation based on cumulant generating function evaluated by using the null distribution, which beats the normal approximation over the entire range.

CHAPTER 5
COMPARISON OF THE TAIL PROBABILITY APPROXIMATION BY THE METHOD OF EDGEWORTH EXPANSION AND SADDLEPOINT APPROXIMATION

Let

$$\alpha(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2} \quad ; \quad -\infty < X < \infty$$

and $D = \frac{d}{dX}$.

Now, consider successive derivatives of $\alpha(X)$ with respect to X. We have

$$\begin{aligned} D\alpha(X) &= -X\cdot\alpha(X) \\ D^2\alpha(X) &= (X^2-1)\cdot\alpha(X) \\ D^3\alpha(X) &= (3X-X^3)\cdot\alpha(X) \\ D^4\alpha(X) &= (X^4-6X^2+3)\cdot\alpha(X) \end{aligned}$$

and so on.

Edgeworth (1904) derived a series to approximate a density function, which for a standardized case (when $k_1 = 0$ and $k_2 = 1$) reduces to,

$$f(X) = \exp\{-k_3\frac{D^3}{3!} + k_4\frac{D^4}{4!} - \ldots.\} \cdot \alpha(X).$$

Edgeworth expansion, even though it has been used to a large extent both in past and present for approximation (of density or distribution), still has problems associated with it. These have been discussed in several books as well as papers. Statistically speaking, it is not as important to use a series to represent a density function as it is to make sure that a finite number of terms can result in a satisfactory approximation. It is important here to note that the series used in Edgeworth expansion

- may give negative frequencies, particularly near the tails, when a finite number of terms are used (Kendal, Stuart and Ord (1987) Vol. 1).
- may behave irregularly in the sense that the sum of k terms may result in a worse fit than the sum of (k - 1) terms.

For many statistical purposes, we are most interested in the tails of distribution, and it is here that shortcomings of the series (Edgeworth expansion) approach appear.

P. J. Bickel, F. Gotze and W. R. Van Zwet (1986) under some general and mild assumptions (similar to the ones assumed in this dissertation) gave an Edgeworth expansion for U-statistics (one sample) of degree two. Their results also assume (without loss of generality) a standardized U-Statistic based on a symmetric kernel among some other technical assumptions.

To be specific, if

$$E\, h(X_1, X_2) = 0 \quad ; \quad E\, h^2(X_1, X_2) < \infty,$$

and define

$$g(X_1) = E\,(h(X_1, X_2) \mid X_1), \quad \psi(X,Y) = h(X,Y) - g(X) - g(Y)$$

and also,

$$\hat{U}_N = (N-1)\sum_{i=1}^{N} g(X_i) \quad , \quad \Delta_N = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \psi(X_i, X_j)$$

$\Rightarrow$

$$U_N = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} h(X_i, X_j) = \hat{U}_N + \Delta_N . \tag{5.1}$$

Furthermore

if $\quad \sigma_N^2 = Var(U_N) \quad ; \quad \sigma_g^2 = E\,g^2(X_1) > 0$

and Let,

$$k_3 = \sigma_g^{-3}\{E\,g^3(X_1) + 3\,E\,g(X_1)\,g(X_2)\,\psi(X_1,X_2)\}$$

$$k_4 = \sigma_g^{-4}\{E\,g^4(X_1) - 3\sigma_g^4 + 12\,E\,g^2(X_1)\,g(X_2)\,\psi(X_1,X_2)$$

$$+ 12\,E\,g(X_1)\,g(X_2)\,\psi(X_1,X_3)\,\psi(X_2,X_3)\} \tag{5.2}$$

then finally under the assumption

$$E\,h^4(X_1,X_2) < \infty ,$$

they show that $k_3 \cdot N^{-1/2}$ and $k_4 \cdot N^{-1}$ are asymptotic expressions with error

$o(N^{-1})$ for the third and fourth cumulants of $\sigma_N^{-1} U_N$. their Edgeworth expansion under this setup is given by

$$F_N(x) = \Phi(x) - \phi(x) \left[\frac{k_3}{6} \cdot N^{-1/2} (x^2 - 1) + \frac{k_4}{24} \cdot N^{-1} (x^3 - 3x) \right.$$

$$\left. + \frac{k_3^2}{72} \cdot N^{-1} (x^5 - 10x^3 + 15x) \right] \qquad (5.3)$$

and the related order of approximation is $o(N^{-1})$.

We further replaced $k_3 \cdot N^{-1/2}$ and $k_4 \cdot N^{-1}$ by the exact third and fourth cumulants of $\sigma_N^{-1} U_N$ in (5.3) to optimize the accuracy of their expansion and hence obtain the Edgeworth expansion as follows:

$$F_N(x) = \Phi(x) - \phi(x) \left[\frac{k_3}{6} (x^2 - 1) + \frac{k_4}{24} \cdot (x^3 - 3x) \right.$$

$$\left. + \frac{k_3^2}{72} (x^5 - 10x^3 + 15x) \right] \qquad (5.4)$$

It should also be noted here that in our examples, the equations (5.3) and (5.4) for the two Edgeworth expansions gave the same approximation to the tail

probability upto the third decimal place, i.e., see Table (5.1). Then equation (5.4) was used to approximate the tail probability for comparison with our saddlepoint approximation of tail probability where the cumulant generating function was in turn approximated by using Hoeffding's decomposition, i.e. equation (3.10) in Chapter 3. We carried out this comparison for one sample Wilcoxon Signed Rank test statistic based on a sample of size 17 under the null hypothesis. Note that under the null hypothesis (since the standardized test statistic is symmetric about zero) for this test $k_3 = 0$.

Saddlepoint based approximation did much better over the whole range of values associated with the test statistic as compared to both Bickel, Gotze and Van-Zwet's Edgeworth expansion of a one sample U-statistic of degree 2, i.e. equation (5.3) and Edgeworth expansion based on exact cumulants, i.e. equation (5.4), see Figures 5.1(A-C).

Table 5.1. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) between Exact (Null) Distribution, Normal Approximation, Edgeworth Expansion and Saddlepoint Approximation

W^+	W^+_{std}	Sp_app	Exact	Edge_ex	Edge_Bk	Normal
35	-1.964530	0.0262973	0.0253	0.0239	0.0239	0.0247
40	-1.727840	0.0449721	0.0443	0.0420	0.0420	0.0420
45	-1.491150	0.0731930	0.0727	0.0693	0.0692	0.0680
50	-1.254460	0.1127630	0.1123	0.1076	0.1075	0.1048
55	-1.017770	0.1648830	0.1645	0.1585	0.1583	0.1544
60	-0.781079	0.2296940	0.2293	0.2221	0.2219	0.2174
65	-0.544388	0.3060150	0.3056	0.2974	0.2973	0.2931
70	-0.307698	0.3913040	0.3910	0.3821	0.3820	0.3792
75	-0.071007	0.4818850	0.4816	0.4724	0.4724	0.4717

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

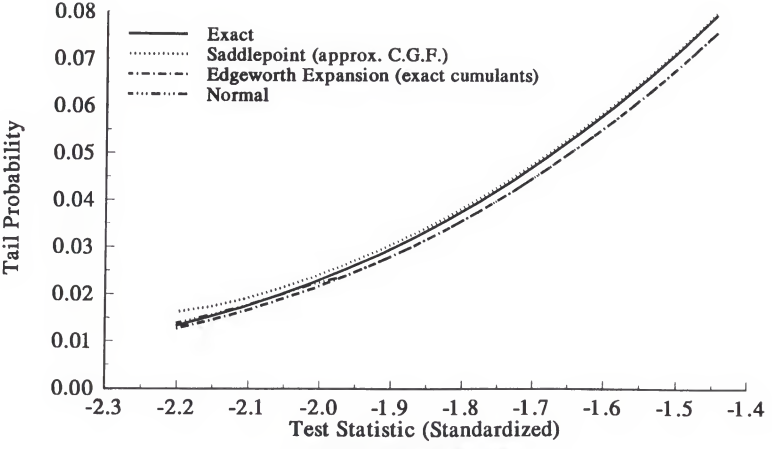
Sp_app = Saddlepoint Approximation Based on Our Approximation

Exact = Null Distribution of Test-Statistic

Edge_ex = Edgeworth Expansion Based on Exact Cumulants (K_1-K_4)

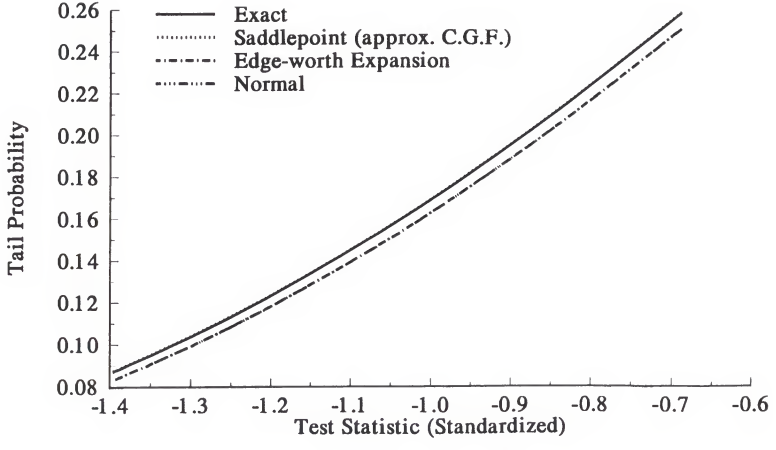
Edge_Bk = Edgeworth Expansion (Bickel,Gotze and Van Zwet)

Normal = Normal Approximation

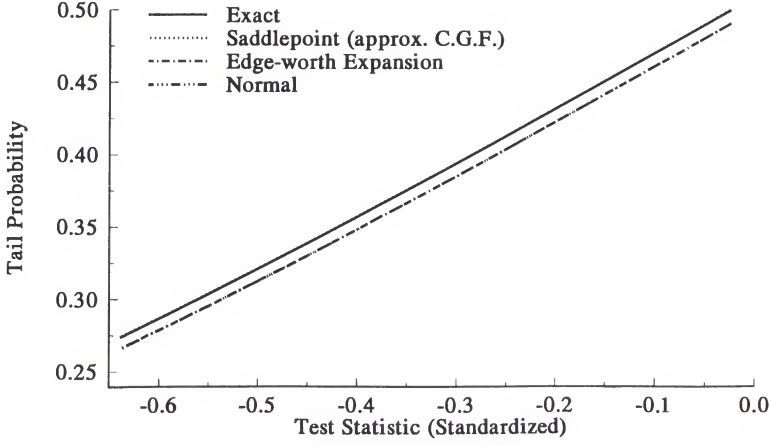


Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddlepoint(approx. C.G.F.), Normal and Edge-worth Expansion

Figure 5.1 (A)



Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddlepoint(approx. C.G.F.), Normal and Edge-worth Expansion
Figure 5.1 (B)



Wilcoxon Signed Rank Test
Comparison of Tail Probability Approximation
Using Exact, Saddlepoint(approx. C.G.F.), Normal and Edge-worth Expansion

Figure 5.1 (C)

CHAPTER 6
POWER OF ONE SAMPLE U-STATISTICS

The computation of the exact power can be quite difficult for the one sample U-statistics even for moderate sample sizes. One of the reasons for this is that the exact distribution of the U-statistics depends on the distribution of the ranks which may not be easy to handle. In view of the asymptotic normality of the U-statistics, it is reasonable to use normal approximation to estimate it's power. It needs to be noted here that the normal approximation may in fact work quite well as long as:

- the sample size is large enough;
- under the alternative hypothesis, the location shift parameter(Δ) is not large;
- the underlying population distribution is not a heavy tailed distribution.

In the absence of any of the above, the normal approximation may not be found to be as accurate as desired. In view of these problems associated with the large sample approximation, we propose that the saddlepoint based our tail probability approximation may provide a viable alternative to normal approximation.

6.1 Discussion and Methods

Observe that the tail probability approximation using the saddlepoint methodology derived under a minimal set of assumptions, in Chapter 3 is a general

result for a one sample U-statistic, based on a kernel of degree K. We also assume that the underlying distribution (population) is a continuous distribution function. Thus, to use equation (2.8) in Chapter 2 to approximate power under an alternative requires us first to approximate the cumulant generating function of the standardized U-statistic (with a kernel of degree K) by equation (3.20) in Chapter 3. To approximate the cumulant generating function, we further need to approximate a few expectations under alternative hypothesis. Standard Numerical Integration techniques can be used to do this. Note, since these expectations are based on functions of kernel of degree K such as,

$$g_1(x_1) = E\,h(X_1, \ldots, X_k \mid X_1 = x_1)$$

$$g_2(x_1,x_2) = E\,h(X_1, \ldots, X_k \mid X_1 = x_1, X_2 = x_2) \qquad (6.1)$$

and,

$$g_3(x_1,x_2,x_3) = E\,h(X_1, \ldots, X_k \mid X_1 = x_1, X_2 = x_2, X_3 = x_3).$$

i.e., some conditional expectations of the kernel. Note that if we are interested only in approximating the power of a U-statistic, then since these conditional expectations of the kernel would be independent of the sample size if the kernel itself is independent of N (sample size), i.e., we would not have to re-evaluate these for different sample sizes for a specific one sample U-statistics.

Occasionally, we will come across examples of non-parametric test statistics

where the kernel may be dependent on the sample size simply because the test statistic is a linear combination of U-statistics (whose tail probability can also be handled by the same methodology). For these we would find a dependency of the above conditional expectations (6.1) on the sample size. One such example is the Wilcoxon Signed Rank test (as discussed in Chapter 3).

In approximating these expectations, care must be taken not to sacrifice the accuracy simply because these expectations are usually quite small, i.e., or order between 10^{-2} and 10^{-4}.

To calculate power under different shift alternatives, we decided to again use one sample Wilcoxon Signed Rank (which is a linear combination of two U-statistics) as an example because of two reasons:

1. It is a linear combination of two U-statistics.
2. The form for $g_1(X_1)$ and $\psi(X_1,X_2)$ has already been specified in Chapter 3.

Note that the tail probability under null hypothesis (p-value) is independent of the underlying distribution (as long as it is continuous and symmetric about 0) since the U-statistics is distribution free. Thus, these expectations should be the same under all the continuous distributions which are symmetric about zero. For one sample Wilcoxon Signed Rank test, we only had to evaluate 4 different expectations which were as follows:

(i) $E_1 = E\, g_1(x)\, g_1(y)\, \psi_2(x,y)$, $\quad (say)$

(ii) $E_2 = E\, g_1^2(x)\, g_1(y)\, \psi_2(x,y)$

(iii) $E_3 = E\, g_1(x)\, g_1(y)\, \psi_2(x,z)\, \psi_2(y,z)$ (6.2)

(iv) $E_4 = E\, \psi_2^2(x,y)$

where

$$g_1(X) = \left(1 - F(-X) - \int (1 - F(-X))\, dF(X)\right) + \frac{1}{N-1}\left[F(0) - 1 + 1_{[0,\infty]}(X)\right]$$

and

$$\psi_2(X, Y) = 1_{[0,\infty]}(X + Y) + F(-X) + F(-Y) + \int (1 - F(-X))\, dF(X)$$

Exact power for Wilcoxon Signed Rank test is not easy to calculate. It has been tabulated by Kotz (1963) for normal shift alternatives and by Arnold (1965) for shifted t-distributions with 1/2, 1, 2, and 4 degrees of freedom for sample sizes between 5 and 10. Usually a large sample approximation (normal) is used to calculate power which also may be inaccurate, as discussed at the beginning of this chapter. There is also a discussion in Lehman's (1975) book "Nonparametrics" about these issues.

If a test is expressed in terms of critical region $W_s \geq c$, its power against any fixed alternative

H_1: *F is symmetric about* $\Delta \neq 0$.

is given by

$$\Pi_{\Delta} = P\,(W_s \geq C) = P\left[\frac{W_s - E_{\Delta}(W_s)}{(Var_{\Delta}(W_s))^{1/2}} \geq \frac{C - E_{\Delta}(W_s)}{(Var_{\Delta}(W_s))^{1/2}}\right] \tag{6.3}$$

where if

$$P_1 = P_{\Delta}(z > 0) \quad , \quad P_2 = P_{\Delta}(z_1 + z_2 > 0)$$

and

$$P_3 = P_{\Delta}(z_1 + z_2 > 0 \text{ and } z_1 + z_3 > 0)$$

then

$$E_{\Delta}(W_s) = \frac{N(N-1)}{2} \cdot P_2 + N P_1 \tag{6.4}$$

and

$$Var_{\Delta}(W_s) = N(N-1)(N-2)(P_3 - P_2^2) + \frac{N(N-1)}{2}\left[2(P_1 - P_2)^2 + 3P_2(1 - P_2)\right] + N P_1(1 - P_1) \tag{6.5}$$

Note that even this normal approximation discussed above requires the computation of P_1, P_2, and P_3. There is another asymptotic approximation of power

discussed by Lehman (1975), but it only works well for very small (delta's) under shift alternatives and large enough sample sizes. According to Lehman this may in fact give completely misleading results in the Cauchy alternatives. Even the normal approximation (eq. 6.3) may not be accurate enough sometimes, especially when there is a heavier tailed distribution, e.g., Cauchy, the sample size is not large enough or delta, the location shift parameter is not small. As a result this approximation may only work better if the underlying distribution is either normal or quite similar to normal distribution, delta is small and the sample size is large enough.

Computation of power of Wilcoxon Signed Rank test for comparison between Saddlepoint method using the approximated cumulant generating function, normal approximation (equation 6.3) and exact was performed under normal, logistic and Cauchy alternatives. These three distributions were chosen simply to illustrate how well the tail probability approximation based on the saddlepoint method using the approximated cumulant generating function can handle the tail's getting thicker, i.e., normal density goes to zero at a very fast rate of $\exp(-x^2)$ while the Cauchy density goes to zero at a relatively much slower rate of x^{-2}. On the other hand for the logistic distribution, the density goes to zero at a rate of e^{-x} which is between that of normal and Cauchy. Since only shift alternatives were going to be considered,thus for consistency between the comparisons for standard, normal, logistic and Cauchy distributions (all symmetric about zero), alternatives were symmetric about their 60, 70 and 80 percentiles in each of the three cases. As a result, nine different alternatives were taken into consideration, three for each distribution.

The expectations needed to approximate the cumulant generating function further needed to approximate the power by the saddlepoint method have been tabulated in Tables 6.1-6.3 on next two pages. Note that there are only four expectations needed to successfully use our saddlepoint based approximation to approximate the tail probability for the Wilcoxon Signed Rank test under different alternatives.

6.2 Results

We used the Wilcoxon Signed Rank (N=17) test (same as in 3rd and 4th chapters) as our example for comparison purposes. As per our discussion in the previous section, we considered location shift alternatives under normal, logistic and Cauchy distributions. Standardized forms of these distributions (each symmetric about zero) were used under the null hypothesis for the test statistic. While for location shift alternatives, we used 60th, 70th and 80th percentiles of these standardized distributions as values for delta (Δ), the location shift parameter.

Nine tables (6.4 - 6.12) have been produced to compare the exact tail probabilities under the various location shift alternatives for each of the three underlying distributions with both our saddlepoint based tail probability approximation and large sample (normal) approximation. We plotted the left tail probabilities for the 80th percentile location shift alternatives under both normal (Figure 6.1) and Cauchy (Figure 6.2) distributions. This was done simply to get a flavor for these results. All the exact tail probabilities were calculated by 10 million

Table 6.1. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Power Calculation of Wilcoxon Signed Rank Test under Normal Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	-0.00374194	-0.00480049	-0.00291832
E_2	-0.00243942	-0.000533953	0.000534172
E_3	0.00164375	0.00125209	0.00056999
E_4	0.0801455	0.0695841	0.0497001

Where

$$E_1 = E\, g_1(x)\, g_1(y)\, \psi_2(x,y)$$

$$E_2 = E\, g_1^2(x)\, g_1(y)\, \psi_2(x,y)$$

$$E_3 = E\, g_1(x)\, g_1(y)\, \psi_2(x,z)\, \psi_2(y,z)$$

$$E_4 = E\, \psi_2^2(x,y)$$

Table 6.2. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Power Calculation of Wilcoxon Signed Rank Test under Logistic Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	-0.00241817	-0.00331032	-0.00220906
E_2	-0.00268175	-0.00106512	-0.00000398
E_3	0.00158255	0.0011335	0.0005165
E_4	0.0797383	0.0685381	0.0489896

Note: E_1 - E_4 are defined the same as on page 110.

Table 6.3. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Power Calculation of Wilcoxon Signed Rank Test under Cauchy Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	0.00189138	0.00301712	0.00283385
E_2	-0.00346422	-0.00331040	-0.00266631
E_3	0.00156578	0.00111967	0.00055827
E_4	0.07843870	0.06405560	0.04215510

Note: E_1 - E_4 are defined the same as on page 110.

simulations. The mean and variance under each of these alternatives were calculated by equations 6.4 and 6.5.

It should be noted that our saddlepoint based tail probability approximation does consistently better than normal approximation over each of the nine alternatives for this example. Again for an extreme test statistic value, when the test statistic is quite large, the saddlepoint is also relatively larger and as a result our saddlepoint approximation (calculated by using the cumulant generating function based on Hoeffding's decomposition) may tend to lose it's accuracy. The main reason for this may be that the saddlepoint itself may not be too accurate since the cumulant generating function approximated by Hoeffding's decomposition is not accurate enough. If this is a problem then we suggest to include a few more higher order terms to improve its accuracy.

Table 6.4. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Normal Alternative(Δ=0.253347) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
33	-3.1914	0.0012	0.0010	0.0007
38	-2.9429	0.0027	0.0026	0.0016
43	-2.6945	0.0055	0.0053	0.0035
48	-2.4461	0.0106	0.0103	0.0072
53	-2.1976	0.0191	0.0188	0.0149
58	-1.9492	0.0328	0.0324	0.0256
63	-1.7008	0.0536	0.0533	0.0444
68	-1.4524	0.0835	0.0834	0.0732
73	-1.2039	0.1247	0.1247	0.1143
78	-0.9555	0.1785	0.1789	0.1696
83	-0.7071	0.2454	0.2462	0.2497
88	-0.4587	0.3248	0.3257	0.3232
93	-0.2102	0.4138	0.4151	0.4168
Right Tail				
98	0.0382	0.5099	0.5093	0.4848
103	0.2866	0.4132	0.4128	0.3872
108	0.5351	0.3194	0.3194	0.2963
113	0.7835	0.2340	0.2345	0.2167
118	1.0319	0.1611	0.1624	0.1510
123	1.2803	0.1030	0.1057	0.1002
128	1.5288	0.0603	0.0676	0.0632

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.5. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Normal Alternative(Δ=0.524401) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
53	-3.7081	0.0008	0.0008	0.0001
58	-3.4172	0.0018	0.0018	0.0002
63	-3.1263	0.0036	0.0037	0.0008
68	-2.8354	0.0069	0.0065	0.0023
73	-2.5445	0.0127	0.0123	0.0054
78	-2.2536	0.0226	0.0222	0.0120
83	-1.9627	0.0383	0.0382	0.0248
88	-1.6718	0.0626	0.0628	0.0472
93	-1.3809	0.0971	0.0990	0.0826
98	-1.0900	0.1479	0.1495	0.1378
103	-0.7991	0.2141	0.2162	0.2120
108	-0.5082	0.2974	0.3000	0.3057
113	-0.2173	0.3966	0.4000	0.4141
Right Tail				
118	0.0737	0.5154	0.5371	0.4706
123	0.3646	0.4002	0.3981	0.3577
128	0.6555	0.2874	0.2845	0.2561
133	0.9464	0.1868	0.1820	0.1719
138	1.2373	0.1006	0.0991	0.1080
143	1.5282	0.0497	0.0437	0.0632
148	1.8191	0.0170	0.0163	0.0345

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.6. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Normal Alternative(Δ=0.841621) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
83	-4.0576	0.0013	0.0008	0
88	-3.6574	0.0026	0.0020	0.0001
93	-3.2572	0.0055	0.0047	0.0005
98	-2.8569	0.0110	0.0103	0.0021
103	-2.4567	0.0214	0.0210	0.0070
108	-2.0565	0.0398	0.0403	0.0198
113	-1.6562	0.0707	0.0727	0.0488
118	-1.2560	0.1206	0.1240	0.1045
123	-0.8558	0.1958	0.2001	0.1961
128	-0.4555	0.3016	0.3061	0.3243
133	-0.0553	0.4392	0.4438	0.4779
Right Tail				
138	0.3449	0.4336	0.4224	0.3650
143	0.7451	0.2686	0.2334	0.2280
148	1.1454	0.0986	0.0969	0.1260
153	1.5456	0.0224	0.0377	0.0610

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.7. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Logistic Alternative(Δ=0.405465) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
33	-3.1173	0.0014	0.0014	0.0009
38	-2.8706	0.0030	0.0030	0.0020
43	-2.6239	0.0063	0.0061	0.0043
48	-2.3773	0.0120	0.0117	0.0087
53	-2.1306	0.0216	0.0211	0.0166
58	-1.8839	0.0368	0.0363	0.0298
63	-1.6373	0.0596	0.0591	0.0508
68	-1.3906	0.0923	0.0918	0.0822
73	-1.1439	0.1366	0.1363	0.1263
78	-0.8973	0.1937	0.1937	0.1848
83	-0.6506	0.2637	0.2640	0.2577
88	-0.4039	0.3453	0.3464	0.3531
93	-0.1572	0.4358	0.4445	0.4375
Right Tail				
98	0.0894	0.4882		0.4644
103	0.3361	0.3925	0.3929	0.3684
108	0.5828	0.3013	0.3015	0.2800
113	0.8294	0.2192	0.2201	0.2034
118	1.0761	0.1499	0.1517	0.1409
123	1.3228	0.0953	0.0980	0.0930
128	1.5694	0.0556	0.0644	0.0583

where

W^+ = Wilcoxon Signed Rank Test Statistic
W^+_{std} = Standardized Test-Statistic
Exact = Exact(Simulated) Distribution of T.Stat under the Alternative
Sp_app = Saddlepoint Approximation Based on Our Approximation
Normal = Normal Approximation

Table 6.8. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Logistic Alternative(Δ=0.847298) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
48	-3.7809	0.0005	0.0004	0.0001
53	-3.4987	0.0011	0.0010	0.0002
58	-3.2164	0.0024	0.0022	0.0006
63	-2.9342	0.0048	0.0045	0.0017
68	-2.6519	0.0092	0.0088	0.0040
73	-2.3697	0.0169	0.0163	0.0089
78	-2.0874	0.0294	0.0288	0.0184
83	-1.8052	0.0490	0.0485	0.0365
88	-1.5229	0.0785	0.0780	0.0639
93	-1.2407	0.1201	0.1202	0.1074
98	-0.9584	0.1768	0.1773	0.1689
103	-0.6762	0.2501	0.2507	0.2495
108	-0.3939	0.3382	0.3401	0.3468
113	-0.1117	0.4398	0.4624	0.4555
Right Tail				
118	0.1706	0.4728	0.4773	0.4323
123	0.4528	0.3619	0.3614	0.3253
128	0.7351	0.2569	0.2569	0.2311
133	1.0173	0.1659	0.1670	0.1545
138	1.2996	0.0941	0.0987	0.0969

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.9. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Logistic Alternative(Δ=1.38629) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
78	-3.9934	0.0011	0.0008	0.0000
83	-3.6226	0.0023	0.0019	0.0001
88	-3.2518	0.0047	0.0041	0.0006
93	-2.8811	0.0094	0.0086	0.0020
98	-2.5103	0.0181	0.0172	0.0060
103	-2.1395	0.0333	0.0327	0.0162
108	-1.7688	0.0591	0.0588	0.0385
113	-1.3980	0.0999	0.1005	0.0811
118	-1.0272	0.1612	0.1630	0.1522
123	-0.6565	0.2478	0.2506	0.2558
128	-0.2857	0.3604	0.3658	0.3875
Right Tail				
133	0.0851	0.5321	-	0.4661
138	0.4558	0.3824	0.3782	0.3243
143	0.8266	0.2379	0.2218	0.2042
148	1.1974	0.1138	0.0897	0.1156

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.10. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Cauchy Alternative (Δ=0.32492) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
38	-2.5893	0.0055	0.0058	0.0048
43	-2.3487	0.0108	0.0110	0.0094
48	-2.1080	0.0203	0.0203	0.0175
53	-1.8674	0.0354	0.0351	0.0308
58	-1.6267	0.0585	0.0578	0.0519
63	-1.3861	0.0913	0.0905	0.0829
68	-1.1455	0.1357	0.1351	0.1260
73	-0.9048	0.1928	0.1925	0.1828
78	-0.6642	0.2625	0.2625	0.2533
83	-0.4235	0.3432	0.3435	0.3360
88	-0.1829	0.4320	0.4325	0.4274
Right Tail				
93	0.0578	0.4938	0.4938	0.4770
98	0.2984	0.4009	0.4012	0.3827
103	0.5390	0.3124	0.3132	0.2949
108	0.7797	0.2326	0.2339	0.2180
113	1.0203	0.1645	0.1663	0.1538
118	1.2610	0.1097	0.1119	0.1037
123	1.5016	0.0683	0.0712	0.0666
128	1.7422	0.0395	0.0414	0.0407

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.11. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Cauchy Alternative (Δ=0.726543) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
53	-2.7499	0.0048	0.0050	0.0030
58	-2.4958	0.0093	0.0095	0.0063
63	-2.2416	0.0173	0.0172	0.0125
68	-1.9875	0.0305	0.0299	0.0238
73	-1.7333	0.0508	0.0497	0.0415
78	-1.4792	0.0806	0.0790	0.0695
83	-1.2250	0.1220	0.1200	0.1103
88	-0.9709	0.1762	0.1745	0.1658
93	-0.7167	0.2441	0.2430	0.2368
98	-0.4626	0.3246	0.3244	0.3218
103	-0.2084	0.4147	0.4159	0.4175
Right Tail				
108	0.0457	0.5073	0.5072	0.4818
113	0.2999	0.4092	0.4096	0.3821
118	0.5541	0.3155	0.3161	0.2898
123	0.8022	0.2272	0.2319	0.2095
128	1.0624	0.1548	0.1612	0.1440
133	1.3165	0.0982	0.1093	0.0940

where

W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation

Table 6.12. Comparison of Tail Probability Approximation of Wilcoxon Signed Rank Test Statistic (N=17) under Cauchy Alternative (Δ=1.37638) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W^+	W^+_{std}	Exact	Sp_app	Normal
Left Tail				
78	-2.5280	0.0117	0.0116	0.0057
83	-2.2433	0.0216	0.0209	0.0124
88	-1.9587	0.0370	0.0362	0.0251
93	-1.6740	0.0622	0.0599	0.0471
98	-1.3893	0.0978	0.0952	0.0824
103	-1.1046	0.1457	0.1448	0.1347
108	-0.8199	0.2154	0.2108	0.2061
113	-0.5352	0.2932	0.2935	0.2963
118	-0.2505	0.3845	0.3905	0.4011
Right Tail				
123	0.0342	0.5192	0.5283	0.4864
128	0.3189	0.4165	0.4176	0.3749
133	0.6035	0.3191	0.3146	0.2731
138	0.8882	0.1955	0.2305	0.1872

where

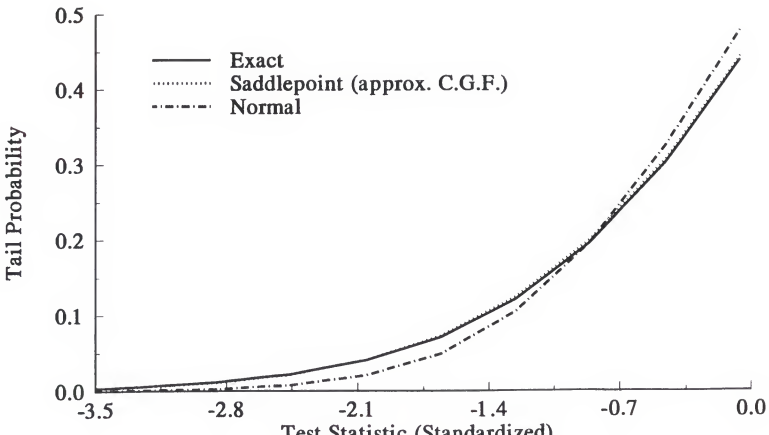
W^+ = Wilcoxon Signed Rank Test Statistic

W^+_{std} = Standardized Test-Statistic

Exact = Exact(Simulated) Distribution of T.Stat under the Alternative

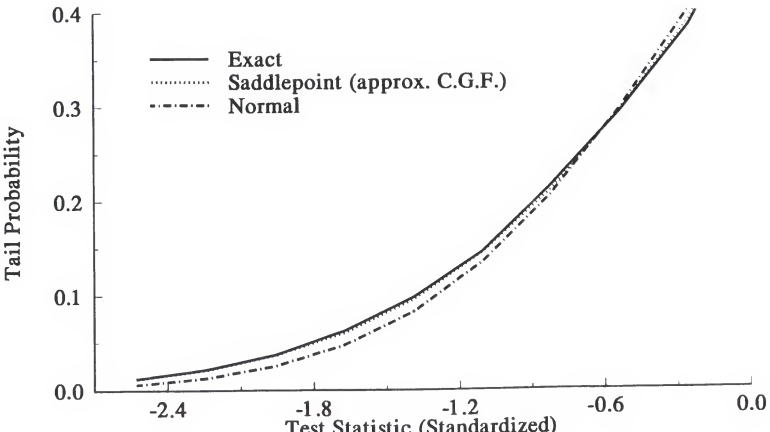
Sp_app = Saddlepoint Approximation Based on Our Approximation

Normal = Normal Approximation



Wilcoxon Signed Rank Test
Comparison of Left Tail Probability Approximation
Under Normal Alternative (Δ=0.841621)
Using Exact, Saddle-Point and Normal Distribution

Figure 6.1



Wilcoxon Signed Rank Test
Comparison of Left Tail Probability Approximation
Under Cauchy Alternative (Δ=1.37638)
Using Exact, Saddle-Point and Normal Distribution

Figure 6.2

CHAPTER 7
POWER: TWO SAMPLE U-STATISTICS

The computation of exact power may be quite tedious and difficult, simply because the distribution of ranks under that alternative may not always be easy to handle. As a result, computation for exact power is usually not a feasible approach to estimation of power. Simulation is a possible alternative; however, it is sensitive to random number generators. A consequence of random fluctuations may cause strange results for tail probability approximation under the alternative.

On the other hand, similar to one sample situation, normal approximation is possibly another alternative available for all two sample U-statistics. Normal approximation works quite well as long as the sample size is large enough, underlying distribution is not a heavy tailed distribution and location change is not too large under the alternative. In the absence of any of the above three conditions, normal approximation may not perform very well. In light of these problems associated with the large sample approximation, we suggest that our saddlepoint approximation may give a reasonable alternative for the tail probability approximation for a two sample U-statistics under various alternatives.

7.1 Discussion and Methods

Our approach, based on saddlepoint approximation in approximating the tail

probability as discussed in Chapter 4, is again quite general since we have tried to derive the tail probability approximation for the two sample U-statistics based on a kernel of degree (K_1,K_2) under very minimal assumptions (including a continuous distribution) associated with the underlying distributions.

Again, to approximate power, we first need to approximate a few expectations under the alternative hypothesis. Also note that since these expectations are really conditional expectations of the kernel (of degree K_1 and K_2), they may be independent of the sample sizes involved if the kernel is independent of the sample sizes. If these expectations are independent of the sample sizes, then we only need to approximate them once under a specific distribution. Once they have been estimated, we can use them for various power calculations of the U-statistic, under the same alternative, but for different sample sizes. Their structure is similar to (6.1) in the previous chapter.

To illustrate power computations for two sample, again we used Mann-Whitney-Wilcoxon test statistic. The form for these expectations has already been specified in Chapter 4 (two sample U-statistics). Again, if we are interested in testing if there is no difference between the underlying population distributions corresponding to the two samples, then the null distribution is simply distribution free, i.e., the p-value is independent of the underlying distribution. As a result the expectations should all be same for each one of the underlying distributions (continuous).

For Mann-Whitney-Wilcoxon Rank Sum test we had to evaluate six different

expectations which were as follows:

$$\text{(i)}\quad E_1 = E\, g_{1,0}(x)\, g_{0,1}(y)\, \psi_{1,1}(x, y)$$

$$\text{(ii)}\quad E_2 = E\, g_{1,0}^2(x)\, g_{0,1}(y)\, \psi_{1,1}(x, y)$$

$$\text{(iii)}\quad E_3 = E\, g_{1,0}(x)\, g_{0,1}^2(y)\, \psi_{1,1}(x, y) \qquad (7.1)$$

$$\text{(iv)}\quad E_4 = E\, \psi_{1,1}^2(x, y)$$

$$\text{(v)}\quad E_5 = E\, g_{1,0}(x)\, g_{1,0}(x')\, \psi_{1,1}(x, y)\, \psi_{1,1}(x', y)$$

$$\text{(vi)}\quad E_6 = E\, g_{0,1}(y)\, g_{0,1}(y')\, \psi_{1,1}(x, y)\, \psi_{1,1}(x, y').$$

Where, under the location shift model we have,

$X_1, \ldots, X_m \sim F(X)$, symmetric, about 0.

$Y_1, \ldots, Y_n \sim G(Y) = F(Y - \Delta)$, symmetric about $\Delta \neq 0$.

if $p = E\, 1_{[0,\infty]}(Y - X)$

then

$$g_{1,0}(X) = 1 - G(X) - p = (1 - p) - G(X)$$

$$g_{0,1}(Y) = F(Y) - p \qquad (7.2)$$

and

$$\psi_{1,1}(X, Y) = 1_{[0,\infty]}(Y - X) + p - F(Y) + G(X).$$

The computation of exact distribution (based on ranks) for two sample Mann-

Whitney statistics under alternatives to hypothesis F = G can be quite tedious and difficult to calculate under different kinds of location shift alternatives. Milton (1970) provides the probabilities (tail) of all $\binom{n+m}{n}$ possible rankings of X's and Y's for $1 \leq n \leq m \leq 7$ for normal shift alternatives and has used these results to obtain exact values of the power of the Wilcoxon test for these sample sizes. Bell, Moser, and Thompson (1966) give the power of the Wilcoxon test for rectangular shift alternatives while Haynam and Govindarajulu (1966) do so for exponential shift alternatives for some combinations of m and n. For double exponential shift alternatives, sample sizes m = n = 5, and $\alpha = 0.05$, Ramsey (1971) has also obtained power for this test.

Usually, a normal approximation is used to approximate power for Mann-Whitney-Wilcoxon test. This has also been discussed in Lehmann's book (1975) on nonparametrics. This approximation is essentially based on the fact that

$$\frac{W_{xy} - E(W_{xy})}{\sqrt{Var(W_{xy})}} \rightarrow N(0, 1) \tag{7.3}$$

as m,n tend to infinity for any fixed distributions F and G. Here W_{xy} is Mann-Whitney form of the test statistic. For this, we explicitly need to calculate mean and variance of W_{xy}. For this, we further need to estimate

$$p_1 = P(X < Y),\ p_2 = P(X < Y \text{ and } X < Y')$$

and

$$p_3 = P (X < Y \text{ and } X' < Y).$$

Then

$$E (W_{xy}) = m \cdot n \cdot p_1$$

and

$$Var (W_{xy}) = m\, n\, p_1 (1 - p_1) + m\, n (n - 1) (p_2 - p_1^2) + n\, m\, (m - 1)(p_3 - p_1^2).$$

Also observe that $p_2 = p_3$ in the shift model provided the distribution is symmetric. It should be noted that this normal approximation works well for small values of Δ. As Δ becomes large, the distribution of W_{xy} becomes less symmetric and, therefore, normal approximation less accurate. Lehmann states in his book that not much is known about non-normal distributions. But, it is believed that for these distributions, this approximation may be accurate for very large sample sizes.

Again, comparisons were carried out for power of Mann-Whitney-Wilcoxon test between saddlepoint method using approximated cumulant generating function, normal approximation (equation 7.3) and exact were performed under normal, logistic, and Cauchy shift alternatives. Similar to one sample case, the location shift under the alternative was 60, 70 and 80 percentiles from the standardized distributions in each of the three (normal, logistic and Cauchy) cases. Thus, in all nine different alternatives were taken into consideration.

The expectations (E_1 ... E_6) needed to approximate the cumulant generating

Table 7.1. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Normal Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	-0.00170874	-0.00312077	-0.00387317
E_2	-0.00255211	-0.00191251	-0.000987123
E_3	-0.00255211	-0.00191251	-0.000987123
E_4	0.082541	0.0799171	0.0744752
E_5	0.00137143	0.00130456	0.00114007
E_6	0.00137143	0.00130456	0.00114007

Where

$$E_1 = E\, g_{1,0}(x)\, g_{0,1}(y)\, \psi_{1,1}(x, y)$$

$$E_2 = E\, g_{1,0}^2(x)\, g_{0,1}(y)\, \psi_{1,1}(x, y)$$

$$E_3 = E\, g_{1,0}(x)\, g_{0,1}^2(y)\, \psi_{1,1}(x, y)$$

$$E_4 = E\, \psi_{1,1}^2(x, y)$$

$$E_5 = E\, g_{1,0}(x)\, g_{1,0}(x')\, \psi_{1,1}(x, y)\, \psi_{1,1}(x', y)$$

$$E_6 = E\, g_{0,1}(y)\, g_{0,1}(y')\, \psi_{1,1}(x, y)\, \psi_{1,1}(x, y').$$

Table 7.2. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Logistic Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	-0.00108731	-0.00202017	-0.00256961
E_2	-0.00261129	-0.0021117	-0.00131622
E_3	-0.00261047	-0.00211013	-0.00131369
E_4	0.0824082	0.0793982	0.0731502
E_5	0.00135642	0.00124981	0.00103606
E_6	0.00135642	0.00124981	0.00103606

Note: E_1 - E_6 are defined the same as on page 131.

Table 7.3. Computation of Expectations Needed to Approximate the Cumulant Generating Function for Cauchy Shift Alternatives

	60 Percentile	70 Percentile	80 Percentile
E_1	0.000905932	0.00183317	0.0026385
E_2	-0.00277992	-0.00277597	-0.0026872
E_3	-0.00277992	-0.00277597	-0.0026872
E_4	0.0820116	0.0772355	0.0655041
E_5	0.00135368	0.00122829	0.000940184
E_6	0.00135368	0.001220829	0.000940184

Note: E_1 - E_6 are defined the same as on page 131.

function under the different shift alternatives, have been tabulated in Tables 7.1 for normal, 7.2 for logistic, and 7.3 for Cauchy.

7.2 Results

Wilcoxon Rank Sum statistic in it's Mann-Whitney form (n=5,m=10) was used as an example for comparison purposes. We again considered location shift alternatives under normal, logistic and Cauchy distributions. Standardized forms of these distributions (each symmetric about zero) were used as a underlying population distribution for one population while for the other, a corresponding location shift alternative was used as the population distribution. The location shift parameter, delta (Δ) took, 60^{th}, 70^{th} and 80^{th} percentiles of the corresponding standardized distribution as it's values.

Nine tables (7.4 - 7.12) help us to compare the exact tail probability of the Wilcoxn-Mann-Whitney test statistic with both, the proposed saddlepoint based tail probability approximation and the large sample (normal) approximation. To get a better feel for these results we plotted left tail probabilities (see Figures 7.1 and 7.2) given in Tables 7.8 (logistic with $\Delta = 0.847298$) and 7.11 (Cauchy with $\Delta = 0.726543$) for the exact and the two approximations. Again we used 10 million simulations to compute the exact tail probability in these tables. Equations 7.4 and 7.5 were used to calculate the mean and variance under each of the nine alternatives.

Again, our proposed saddlepoint based tail probability performed consistently better as compared to normal approximation under all the alternatives. If it is felt

Table 7.4. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Normal Alternative(Δ=0.253347) between Exact Distribution, Normal approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
9	-2.4249	0.0098	0.0088	0.0077
11	-2.1769	0.0190	0.0178	0.0147
13	-1.9289	0.0343	0.0330	0.0269
15	-1.6808	0.0577	0.0567	0.0464
17	-1.4328	0.0913	0.0910	0.0760
19	-1.1848	0.1368	0.1377	0.1180
21	-0.9368	0.1956	0.1974	0.1744
23	-0.6888	0.2672	0.2697	0.2455
25	-0.4408	0.3500	0.3528	0.3297
27	-0.1928	0.4410	0.4435	0.4236
Right Tail				
29	0.0553	0.5116	0.5122	0.4780
31	0.3033	0.4162	0.4175	0.3808
33	0.5513	0.3244	0.3266	0.2907
35	0.7993	0.2406	0.2436	0.2121
37	1.0473	0.1687	0.1720	0.1475
39	1.2953	0.1107	0.1146	0.0976
41	1.5433	0.0671	0.0739	0.0614

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.5. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Normal Alternative(Δ=0.524401) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
9	-3.0007	0.0026	0.0022	0.0013
11	-2.7423	0.0057	0.0049	0.0031
13	-2.4840	0.0112	0.0100	0.0065
15	-2.2257	0.0205	0.0191	0.0130
17	-1.9673	0.0356	0.0341	0.0246
19	-1.7090	0.0585	0.0574	0.0437
21	-1.4506	0.0914	0.0911	0.0734
23	-1.1923	0.1365	0.1371	0.1166
25	-0.9340	0.1948	0.1967	0.1752
27	-0.6756	0.2667	0.2698	0.2496
29	-0.4173	0.3512	0.3548	0.3382
31	-0.1589	0.4454	0.4505	0.4369
Right Tail				
33	0.0994	0.5049	0.5134	0.4604
35	0.3577	0.4050	0.4064	0.3603
37	0.6161	0.3086	0.3111	0.2689
39	0.8744	0.2206	0.2248	0.1909
41	1.1328	0.1462	0.1526	0.1287

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.6. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic(n=5,m=10) under Normal Alternative(Δ=0.841621) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
11	-3.5396	0.0011	0.0008	0.0002
13	-3.2588	0.0023	0.0018	0.0006
15	-2.9780	0.0048	0.0040	0.0015
17	-2.6970	0.0092	0.0081	0.0035
19	-2.4162	0.0170	0.0155	0.0078
21	-2.1353	0.0296	0.0281	0.0164
23	-1.8545	0.0493	0.0480	0.0318
25	-1.5736	0.0783	0.0780	0.0578
27	-1.2928	0.1196	0.1204	0.0980
29	-1.0119	0.1750	0.1774	0.1558
31	-0.7311	0.2461	0.2500	0.2324
33	-0.4502	0.3323	0.3374	0.3263
35	-0.1693	0.4315	0.4367	0.4328
Right Tail				
37	0.1115	0.5149	0.5131	0.4556
39	0.3924	0.4048	0.4048	0.3474
41	0.6732	0.2970	0.2989	0.2504
43	0.9541	0.1992	0.2027	0.1700
45	1.2349	0.1195	0.1240	0.1084

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.7. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Logistic Alternative(Δ=0.405465) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
7	-2.6440	0.0048	0.0047	0.0040
9	-2.3964	0.0103	0.0097	0.0083
11	-2.1489	0.0200	0.0189	0.0158
13	-1.9013	0.0360	0.0349	0.0286
15	-1.6538	0.0604	0.0597	0.0491
17	-1.4062	0.0952	0.0955	0.0798
19	-1.1586	0.1424	0.1439	0.1233
21	-0.9111	0.2027	0.2052	0.1811
23	-0.6635	0.2756	0.2789	0.2535
25	-0.4160	0.3594	0.3629	0.3387
27	-0.1684	0.4509	0.4524	0.4331
Right Tail				
29	0.0792	0.5016	0.5113	0.4780
31	0.3267	0.4065	0.4080	0.3808
33	0.5743	0.3156	0.3180	0.2907
35	0.8218	0.2335	0.2364	0.2121
37	1.0694	0.1630	0.1664	0.1475
39	1.3170	0.1067	0.1105	0.0976
41	1.5645	0.0644	0.0709	0.0614

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.8. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic(n=5,m=10) under Logistic Alternative(Δ=0.847298) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
9	-2.9362	0.0029	0.0023	0.0016
11	-2.6797	0.0063	0.0053	0.0037
13	-2.4232	0.0124	0.0109	0.0077
15	-2.1668	0.0227	0.0209	0.0151
17	-1.9103	0.0393	0.0374	0.0280
19	-1.6538	0.0640	0.0626	0.0491
21	-1.3973	0.0994	0.0989	0.0812
23	-1.1408	0.1469	0.1479	0.1270
25	-0.8844	0.2079	0.2104	0.1882
27	-0.6279	0.2823	0.2860	0.2650
29	-0.3714	0.3686	0.3727	0.3552
Right Tail				
33	0.1416	0.4869	0.4885	0.4437
35	0.3980	0.3882	0.3908	0.3453
37	0.6545	0.2941	0.2988	0.2564
39	0.9110	0.2092	0.2168	0.1811
41	1.1675	0.1379	0.1515	0.1215

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.9. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Logistic Alternative(Δ=1.38629) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
14	-3.0234	0.0039	0.0032	0.0012
16	-2.7467	0.0077	0.0066	0.0030
18	-2.4701	0.0145	0.0128	0.0067
20	-2.1935	0.0257	0.0237	0.0141
22	-1.9169	0.0434	0.0413	0.0276
24	-1.6403	0.0698	0.0683	0.0505
26	-1.3637	0.1075	0.1073	0.0863
28	-1.0871	0.1589	0.1603	0.1385
30	-0.8105	0.2254	0.2284	0.2088
32	-0.5339	0.3069	0.3113	0.2967
34	-0.2573	0.4015	0.4069	0.3985
Right Tail				
37	0.1577	0.4946	0.4945	0.4373
39	0.4343	0.3865	0.3889	0.3320
41	0.7109	0.2820	0.2894	0.2386
43	0.9875	0.1891	0.2054	0.1617

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.10. Comparison of Tail Probability Approximation of Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Cauchy Alternative(Δ=0.32492) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
8	-2.4061	0.0090	-	0.0081
10	-2.1601	0.0182	0.0221	0.0154
12	-1.9141	0.0336	0.0336	0.0278
14	-1.6682	0.0571	0.0513	0.0476
16	-1.4222	0.0915	0.0837	0.0775
18	-1.1762	0.1381	0.1395	0.1198
20	-0.9302	0.1982	0.2003	0.1761
22	-0.6842	0.2707	0.2736	0.2469
24	-0.4382	0.3539	0.3571	0.3306
26	-0.1923	0.4450	0.4474	0.4238
Right Tail				
28	0.0537	0.5082	0.4915	0.4786
30	0.2997	0.4139	0.4158	0.3822
32	0.5457	0.3237	0.3269	0.2926
34	0.7917	0.2417	0.2460	0.2143
36	1.0377	0.1713	0.1761	0.1497
38	1.2836	0.1144	0.1111	0.0996
40	1.5296	0.0715	0.0725	0.0631

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.11. Comparison of Tail Probability Approximation for Mann-Whitney-Wilcoxon Test Statistic (n=5,m=10) under Cauchy Alternative (Δ = 0.726543) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
10	-2.5676	0.0070	-	0.0051
12	-2.3176	0.0140	0.0133	0.0102
14	-2.0677	0.0257	0.0246	0.0193
16	-1.8178	0.0443	0.0430	0.0346
18	-1.5678	0.0722	0.0711	0.0585
20	-1.3179	0.1112	0.1108	0.0938
22	-1.0679	0.1627	0.1636	0.1428
24	-0.8180	0.2275	0.2298	0.2067
26	-0.5681	0.3049	0.3081	0.2850
28	-0.3181	0.3927	0.3960	0.3752
30	-0.0682	0.4871	0.4894	0.4728
Right Tail				
32	0.1817	0.4655	0.4659	0.4279
34	0.4317	0.3707	0.3730	0.3330
36	0.6816	0.2816	0.2862	0.2477
38	0.9316	0.2020	0.2093	0.1758
40	1.1815	0.1365	0.1461	0.1187
42	1.4314	0.0846	-	0.0762

where

W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation

Table 7.12. Comparison of Tail Probability Approximation for Wilcoxon-Mnn-Whitney Test Statistic (n=5,m=10) under Cauchy Alternative (Δ=1.37638) between Exact Distribution, Normal Approximation and Saddlepoint Approximations.

W	W_{std}	Exact	Sp_app	Normal
Left Tail				
14	-2.6889	0.0069	0.0063	0.0036
16	-2.4278	0.0133	0.0123	0.0076
18	-2.1666	0.0241	0.0225	0.0151
20	-1.9055	0.0412	0.0390	0.0284
22	-1.6443	0.0665	0.0642	0.0501
24	-1.3832	0.1027	0.1007	0.0833
26	-1.1220	0.1514	0.1505	0.1309
28	-0.8609	0.2147	0.2146	0.1946
30	-0.5997	0.2912	0.2924	0.2743
32	-0.3386	0.3785	0.3817	0.3675
34	-0.0775	0.4739	0.4784	0.4691
Right Tail				
36	0.1837	0.4765	0.4753	0.4271
38	0.4448	0.3745	0.3785	0.3282
40	0.7060	0.2813	0.2884	0.2401
42	0.9671	0.1937	0.1900	0.1667
44	1.2283	0.1275	-	0.1097

where

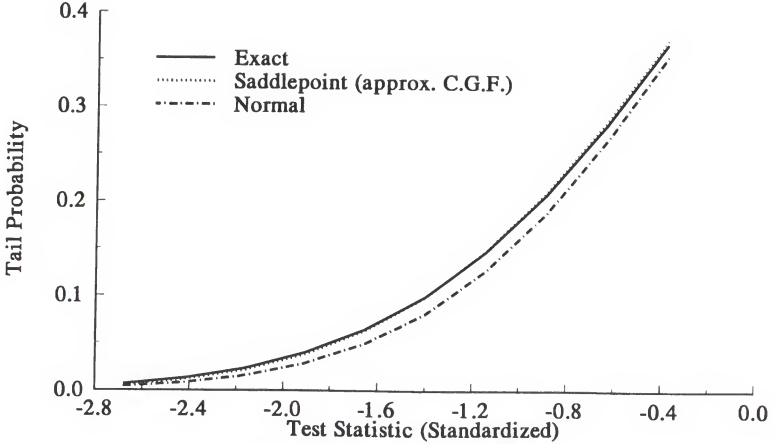
W = Mann-Whitney-Wilcoxon Test Statistic

W_{std} = Standardized Test Statistic

Exact = Exact(Simulated) Distribution of T.Stat Under Alternative

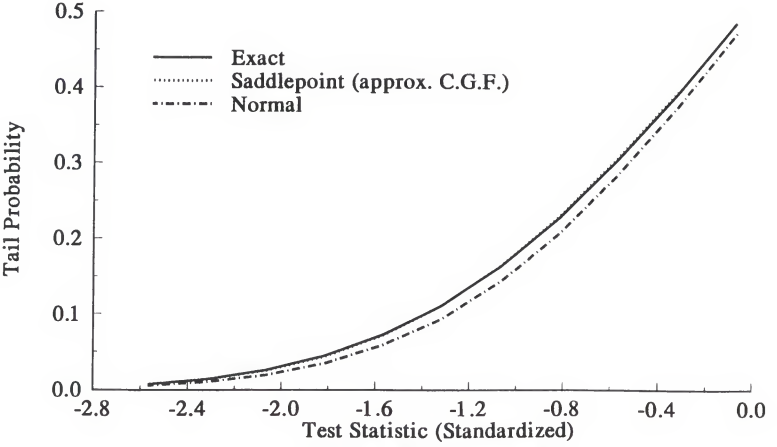
Sp_app = Saddlepoint Approximation Based on Our Result

Normal = Normal Approximation



Wilcoxon-Mann-Whitney Test Statistic
Comparison of Left Tail Probability Approximation
Under Logistic Alternative (Δ=0.847298)
Using Exact, Saddle-Point and Normal Distribution

Figure 7.1



Wilcoxon-Mann-Whitney Test Statistic
Comparison of Left Tail Probability Approximation
Under Cauchy Alternative (Δ=0.726543)
Using Exact, Saddle-Point and Normal Distribution
Figure 7.2

that the saddlepoint approximation to the tail probability is not as accurate as desired (specially in the tails), then it may be worthwhile to consider including more higher order terms in the expansion for the cumulant generating function as given in Chapter 4.

REFERENCES

Arnold, Harvey J. (1965). "Small sample power for the one-sample Wilcoxon test for non-normal shift alternatives." *Annals of Mathematical Statistics* 36, 1767-1778.

Bell, C. B., Moser, J. M. & Thompson, Rory (1966). "Goodness criteria for two-sample distribution-free tests." *Annals of Mathematical Statistics* 37, 133-142.

Bickel, P. J., Gotze, F. & van Zwet, W. R. (1986). "The Edgeworth expansion for *U*-statistics of degree two." *Annals of Statistics* 14, 1463-1484.

Buckle, N., Kraft, C. H. & van Eeden, C. (1969a). "An approximation to the Wilcoxon-Mann-Whitney distribution." *Journal of American Statistical Association* 64, 591-599.

Buckle, N., Kraft, C. H. & van Eeden, C. (1969b). *Tables Prolongées de la Distribution de Wilcoxon-Mann-Whitney,* Presse de l'Université de Montreal, Montreal.

Daniels, H. E. (1954). "Saddlepoint approximations in statistics." *Annals of Mathematical Statististics* 25, 631-650.

Daniels, H. E. (1987). "Tail probability approximations." *International Statistical Review* 55, 37-48.

Dharmadhikari, S. W., Fabian, V. & Jogdeo, K. (1968). "Bounds on the moments of martingales." *Annals of Mathematical Statististics* 39, 1719-1723.

Easton, G. & Ronchetti, E. (1986). "General saddlepoint approximations." *Journal of American Statistical Association* 81, 420-430.

Halmos, P. R. (1946). "The theory of unbiased estimation." *Annals of Mathematical Statististics* 17, 34-43.

Haynam, George E. & Govindarajulu, Z. (1966). "Exact power of Mann-Whitney tests for exponential and rectangular alternatives." *Annals of Mathematical Statististics* 37, 945-953.

Hoeffding, W. (1948). "A class of statistics with asymptotically normal distribution." *Annals of Mathematical Statistics* 19, 293-325.

Hoeffding, W. (1961). "The strong law of large numbers for *U*-statistics." *Institute of Statistics Mimeo Series No 302,* University of North Carolina.

Kendall, M. G., Stuart, A. & Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics,* Vol. 1, 5th Ed., Oxford University Press, New York.

Klotz, Jerome (1963). "Small sample power and efficiency for the one-sample Wilcoxon and normal scores tests." *Annals of Mathematical Statistics* 34, 624-632.

Lehmann, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks,* Holden-Day, Oakland, California.

Lee, A. J. (1990). *U-Statistics,* Marcel Dekker, New York.

Lugannani, R. & Rice, S. (1980). "Saddlepoint approximation for the distribution of the sum of independent random variables." *Advances in Applied Probability* 12, 475-490.

Milton, Roy C. (1970). *Rank Order Probabilities: Two-sample Normal Shift Alternatives,* John Wiley & Sons, New York.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications,* 2nd Ed., Wiley, New York.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics,* Wiley, New York.

Skovgaard, I. M. (1987). "Saddlepoint expansions for conditional distributions." *Journal of Applied Probability* 24, 875-887.

Sundrum, R. M. (1954). "A further Approximation to the distribution of Wilcoxon's statistic in the general case." *Journal of Royal Statistical Society* B16, 255-260.

BIOGRAPHICAL SKETCH

Divakar Sharma was born on July 27, 1963, in Kanpur, India. He was awarded a Bachelor of Science degree in mathematical statistics in 1984, from Hindu College, University of Delhi, India. He also received a Master of Statistics degree in May, 1988, from the University of Florida. He was a resident of Gainesville, Florida, between 1986 and 1992. While working toward his Ph.D. in statistics from the University of Florida, he also worked as a statistical consultant for the Division of Biostatistics, a teaching assistant and an instructor of statistical courses. Since May, 1992 he has been working as a statistician in the Merck Research Laboratories, Blue Bell, Pennsylvania.

Mr. Sharma married Supriya Pal in June, 1990. After graduation he looks forward to doing more consulting and research at Merck Research Laboratories, Blue Bell, Pennsylvania.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Myron N. Chang, Chairman
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it confirms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Malay Ghosh
Professor of Statistics

I certify that I have read this study and that in my opinion it confirms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jonathan J. Shuster
Professor of Statistics

I certify that I have read this study and that in my opinion it confirms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Pejaver V. Rao
Professor of Statistics

I certify that I have read this study and that in my opinion it confirms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Murali Rao
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1993.

_____

Dean, Graduate School